

1 **Select Updates for Recommendations**
2 **for Clinical Laboratory Improvement**
3 **Amendments of 1988 (CLIA) Waiver**
4 **Applications for Manufacturers of**
5 **In Vitro Diagnostic Devices**

7 **Draft Guidance for Industry and**
8 **Food and Drug Administration Staff**

10 ***DRAFT GUIDANCE***

12 **This draft guidance document is being distributed for comment purposes only.**
13 **Document issued on November 29, 2017.**

15 You should submit comments and suggestions regarding this draft document within 60 days of
16 publication in the *Federal Register* of the notice announcing the availability of the draft
17 guidance. Submit electronic comments to <http://www.regulations.gov>. Submit written
18 comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630
19 Fishers Lane, rm. 1061, Rockville, MD 20852. Identify all comments with the docket number
20 listed in the notice of availability that publishes in the *Federal Register*.

22 For questions about this document, please contact the Office of In Vitro Diagnostics and
23 Radiological Health (OIR) at 301-796-5711, Marina Kondratovich, PhD, 301-796-6036 or by
24 email at marina.kondratovich@fda.hhs.gov or Peter Tobin, PhD, 240-402-6169 or by email at
25 peter.tobin@fda.hhs.gov.

27 **When final, this document will update and supersede Section V. of the**
28 **Recommendations for Clinical Laboratory Improvement Amendments of 1988**
29 **(CLIA) Waiver Applications for Manufacturers of In Vitro Diagnostic Devices**
30 **guidance, issued on January 30, 2008.**



33 U.S. Department of Health and Human Services
34 Food and Drug Administration
35 Center for Devices and Radiological Health
Office of In Vitro Diagnostics and Radiological Health

36

Preface

37

38

39

Additional Copies

41

42

43

44

Additional copies are available from the Internet. You may also send an e-mail request to CDRH-Guidance@fda.hhs.gov to receive a copy of the guidance. Please use the document number 16046 to identify the guidance you are requesting.

45

46

47

48

DRAFT

Table of Contents

49

50

51 **I. Introduction and Scope 4**

52 **II. Background and Rationale 5**

53 **III. Revised Section V..... 6**

54 **A. Demonstrating Insignificant Risk of an Erroneous Result – “Accuracy” 6**

55 **1) Quantitative Tests 9**

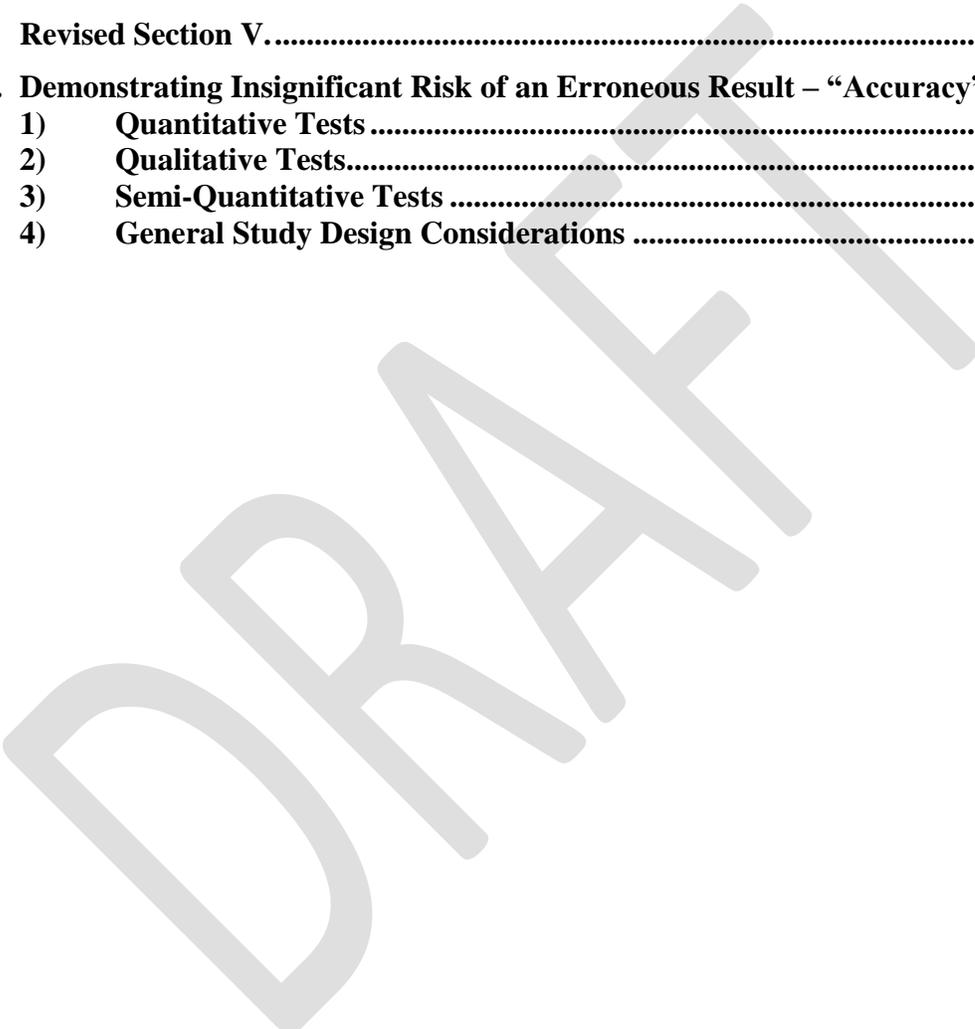
56 **2) Qualitative Tests..... 12**

57 **3) Semi-Quantitative Tests 16**

58 **4) General Study Design Considerations 20**

59

60



61 **Select Updates for Recommendations**
62 **for Clinical Laboratory Improvement**
63 **Amendments of 1988 (CLIA) Waiver**
64 **Applications for Manufacturers of**
65 **In Vitro Diagnostic Devices**

68 **Draft Guidance for Industry and**
69 **Food and Drug Administration Staff**

71 *This draft guidance, when finalized, will represent the current thinking of the Food and Drug*
72 *Administration (FDA or Agency) on this topic. It does not establish any rights for any person*
73 *and is not binding on FDA or the public. You can use an alternative approach if it satisfies*
74 *the requirements of the applicable statutes and regulations. To discuss an alternative*
75 *approach, contact the FDA staff or Office responsible for this guidance as listed on the title*
76 *page.*

78 **I. Introduction and Scope**

79 FDA has developed this draft guidance to implement section 3057 of the 21st Century Cures Act
80 [P.L. 114-255], which requires FDA to revise “Section V. Demonstrating Insignificant Risk of
81 an Erroneous Result — Accuracy” of the guidance “Recommendations for Clinical Laboratory
82 Improvement Amendments of 1988 (CLIA) Waiver Applications for Manufacturers of In Vitro
83 Diagnostic Devices” (“2008 CLIA Waiver Guidance”) that was issued on January 30, 2008. This
84 draft guidance updates FDA’s thinking regarding the appropriate use of comparable performance
85 between a waived user and a moderately complex laboratory user to demonstrate accuracy. The
86 2008 CLIA Waiver Guidance remains in effect, in its current form, until this draft guidance is
87 finalized, at which time the updates in Section III of this draft guidance will supersede the
88 recommendations in Section V. of the 2008 CLIA Waiver Guidance.

89 FDA will incorporate the updates of the final version of this draft guidance into “Section V.
90 Demonstrating Insignificant Risk of an Erroneous Result — Accuracy” of the 2008 CLIA

Contains Nonbinding Recommendations

Draft – Not for Implementation

91 Waiver Guidance. The remainder of the 2008 CLIA Waiver Guidance will not be changed by
92 this update and will remain in effect.

93 For the current edition of the FDA-recognized standard(s) referenced in this document, see the
94 FDA Recognized Consensus Standards Database Web site at
95 <http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfStandards/search.cfm>.

96
97 FDA’s guidance documents, including this guidance, do not establish legally enforceable
98 responsibilities. Instead, guidances describe the Agency’s current thinking on a topic and should
99 be viewed only as recommendations, unless specific regulatory or statutory requirements are
100 cited. The use of the word *should* in Agency guidances means that something is suggested or
101 recommended, but not required.
102

103 **II. Background and Rationale**

104
105 The Secretary of Health and Human Services has delegated to FDA the authority to determine
106 whether particular tests are “simple” and have “an insignificant risk of an erroneous result” under
107 CLIA and thus eligible for waiver categorization (69 FR 22849, April 27, 2004). The Centers
108 for Medicare & Medicaid Services (CMS) is responsible for oversight of clinical
109 laboratories, which includes issuing Certificates of Waiver. CLIA requires that clinical
110 laboratories obtain a certificate before accepting materials derived from the human body for
111 laboratory tests. (42 U.S.C. § 263a(b)). Laboratories that perform only tests that are “simple”
112 and that have an “insignificant risk of an erroneous result” may obtain a certificate of waiver. (42
113 U.S.C. § 263a(d)(2)).
114

115 CLIA, 42 U.S.C. § 263a(d)(3) Examinations and Procedures, as modified by the Food and Drug
116 Administration Modernization Act of 1997 (FDAMA), reads as follows regarding tests that may
117 be performed by laboratories with a Certificate of Waiver:
118

119 The examinations and procedures [that may be performed by a laboratory with a Certificate
120 of Waiver]... are laboratory examinations and procedures that have been approved by the
121 Food and Drug Administration for home use or that, as determined by the Secretary, are
122 simple laboratory examinations and procedures that have an insignificant risk of an erroneous
123 result, including those that — (A) employ methodologies that are so simple and accurate as
124 to render the likelihood of erroneous results by the user negligible, or (B) the Secretary has
125 determined pose no unreasonable risk of harm to the patient if performed incorrectly.
126

127 The 2008 CLIA Waiver Guidance describes recommendations for device manufacturers about
128 study design and analysis for CLIA Waiver by Application to support an FDA determination as
129 to whether the device meets the statutory criteria for waiver described above. This update
130 provides additional details and pathways for demonstrating that a test has an insignificant risk of
131 erroneous result which is a key element for obtaining a CLIA Waiver by Application.
132

Contains Nonbinding Recommendations

Draft – Not for Implementation

133 In developing this specific update, we have considered interactions with stakeholders since the
134 issuance of the final guidance on January 30, 2008.

135 **III. Revised Section V.**

136 **A. Demonstrating Insignificant Risk of an Erroneous Result** 137 **– “Accuracy”**

138
139 As stated previously, a CLIA waiver can be granted for, among others, tests that are “simple
140 laboratory examinations and procedures that have an insignificant risk of an erroneous result”
141 (42 U.S.C. § 263a(d)(3)). This includes tests that employ methodologies that are “so simple and
142 accurate” that the “likelihood of an erroneous result by the user” is rendered “negligible” (42
143 U.S.C. § 263a(d)(3)(A)). One of the key elements for granting a CLIA waiver is that the test is
144 accurate.

145
146 Accuracy is a widely used and generally understood term in clinical laboratory science. For
147 example, the Clinical Laboratory Standard Institute in EP21¹ defines accuracy as “the closeness
148 of an agreement between a test result and the accepted reference value.” Accuracy takes into
149 account the random and systematic components of error. The more accurate a test is, the less
150 error there is, and therefore, the more likely the test is to have an “insignificant risk of erroneous
151 result.” With this in mind, there are various ways that a diagnostic test can be demonstrated to
152 be accurate so that it can be granted a CLIA waiver by application.

153
154 For purposes of this draft guidance and the 2008 CLIA Waiver Guidance, the following terms
155 are defined as:

- 156
- 157 • *Untrained Operator or Waived User*: An operator in waived settings with limited or no
158 training or hands on experience in conducting laboratory testing.
 - 159
 - 160 • *Trained Operator or Moderate Complexity Laboratory User*: A laboratory professional
161 who meets the qualifications to perform moderate complexity testing and with previous
162 training in performing the test.

163
164 Sponsors developing devices designed for the CLIA-waived setting have traditionally taken a
165 sequential route, first obtaining FDA clearance or approval and then submitting data for CLIA
166 waiver determination. In 2012, the MDUFA III Commitment Letter created a single submission
167 pathway referred to as a Dual 510(k) and CLIA Waiver application in which an applicant has an
168 option to apply for marketing authorization and CLIA waiver concurrently within one
169 submission. Proposed recommendations for Dual submissions are provided in the draft guidance
170 [“Recommendations for Dual 510\(k\) and CLIA Waiver by Application Studies”](#)

¹ CLSI. *Evaluation of Total Analytical Error for Quantitative Medical Laboratory Measurement Procedures*. 2nd
ed. CLSI document EP21; Wayne, PA: Clinical and Laboratory Standards Institute, 2016.

Contains Nonbinding Recommendations

Draft – Not for Implementation

171 [https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocu](https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM586502.pdf)
172 [ments/UCM586502.pdf](https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM586502.pdf), which, when finalized, would represent FDA’s current thinking on
173 recommendations for Dual submissions. This draft guidance outlines two options for a
174 sequential route for a CLIA waiver in which safety and effectiveness or substantial equivalence
175 is established first, followed by a separate application containing an evaluation of the test’s
176 accuracy in a CLIA-waived setting:

177
178 **Option 1:** If the sponsor chooses to demonstrate the accuracy of the test when performed
179 by trained operators as part of the marketing submission through comparison to a
180 traceable calibration method (or reference method)², the sponsor can leverage these data
181 in combination with a new study to demonstrate agreement between results of the test
182 performed by untrained operators and trained operators in the waiver by application
183 submission. Valid statistical methods may then be used to estimate the accuracy of the
184 test when performed by untrained operators using a combined analysis of the data from
185 both studies.

186
187 **Option 2:** If the sponsor chooses to demonstrate substantial equivalence or safety and
188 efficacy for the test when performed by trained operators in the marketing application
189 without demonstrating accuracy through a comparison to a traceable calibration method
190 (or reference method), the sponsor should demonstrate accuracy of the test when
191 performed by untrained operators through direct comparison to a traceable calibration
192 method (or reference method), or other comparative method performed in a laboratory
193 setting by trained operators in the waiver application.

194
195 Choosing an option:

196
197 In many cases, a premarket submission for a test contains information about the test performed
198 by laboratory professionals as compared to a reference or traceable calibration method, which
199 allows FDA to assess the test’s performance with trained operators. Under Option 1, a second
200 comparison is then made to assess the level of agreement between the performance of trained
201 and untrained operators. Because Option 1 includes two comparisons in a step-wise fashion, it
202 introduces more potential sources of bias and higher levels of uncertainty (e.g., an increase of
203 confidence intervals, potential bias due to differences in patient populations between studies) in
204 the estimation of the candidate test’s performance in the hands of the untrained operators as
205 compared to Option 2. For quantitative tests, for example, data analysis of the agreement study
206 between test results in the hands of untrained and trained operators in Option 1 provides only an
207 estimation of a difference in systematic errors of the test in the hands of untrained and trained
208 operators at medical decision levels. This systematic difference presents an additional source of
209 error in the accuracy estimates of the test in the hands of untrained operators. However, if a
210 quantitative test’s performance in the hands of trained operators is sufficiently good, there may
211 be sufficient allowable error left to accommodate the additional bias associated with this

² CLSI. *Metrological Traceability and Its Implementation; A Report*. CLSI document EP32-R, Wayne, PA: Clinical and Laboratory Standards Institute, 2006.

Contains Nonbinding Recommendations

Draft – Not for Implementation

212 approach. For qualitative tests, agreement studies (i.e., comparison of the performance of the
213 test by an untrained operator versus the performance of the test by a trained operator) in Option 1
214 typically demonstrate levels of agreement less than 100%, and applicants should consider
215 whether the accuracy demonstrated in the marketing application and the expected agreement are
216 high enough such that a combined statistical analysis will support that the test when performed
217 by untrained operators is accurate. For those sponsors that want the best study to determine
218 accuracy, the study design described in Option 2 is preferred for all types of tests (qualitative,
219 semi-quantitative and quantitative), even when accuracy was determined in the marketing
220 application with trained users.

221
222 In the case of certain tests, a direct estimation of the accuracy of the test in the hands of trained
223 users may not be provided in the pre-market submission. For example, for moderate-risk tests,
224 FDA premarket review evaluates whether the test when performed by laboratory professionals
225 has substantially equivalent performance to that of a predicate device and may not include a
226 comparison to a reference or traceable calibration method. If data are not available to support
227 accuracy of the test with trained operators compared to a reference or traceable calibration
228 method, these tests should use Option 2 for obtaining CLIA waiver by application. Option 1 is
229 generally not appropriate for this scenario because, without some connection to a reference or
230 traceable calibration method or initial understanding of a test's accuracy, it is not possible to
231 determine the likelihood of an erroneous result of the test in the hands of untrained operators by
232 only comparing results of the test in the hands of trained and untrained operators.

233
234 If using Option 1, applicants should provide the following summary information regarding the
235 study, demonstrating the accuracy of the test when performed by trained operators, which they
236 are leveraging:

- 237
- 238 • The number and description of sites in the comparison study, and the number and
239 description of trained operators per individual site.

 - 240 • The number of samples in the study, whether these samples were patient samples (if so,
241 provide inclusion/exclusion criteria), and if some samples were archived or contrived, the
242 number of such samples and explanation of how these samples were selected or prepared.
243 The study subjects/samples should be representative of the intended use population at
244 CLIA waived sites.

 - 245 • The duration of the comparison study and the sources of variability that were included
246 (e.g., different days, different runs, different lots and so on).

 - 247 • The type of comparator method (CM) used in the study (e.g., traceable calibration
248 method, reference method, gold standard, etc.). If the CM is a quantitative traceable
249 calibration method or qualitative traceable cutoff method, provide information supporting
250 traceability of calibrators or cutoffs, as appropriate.

 - 251 • Statistical analysis, depending on the type of assay:

Contains Nonbinding Recommendations

Draft – Not for Implementation

- 252 ○ For quantitative assays, provide the linear regression equation for the test results
253 with trained operators (Y-axis) versus CM results with trained operators (X-axis)
254 as $\text{Test}_{\text{Trained}} = A_1 \cdot \text{CM} + A_0$, the 95% confidence intervals for the slope A_1 and the
255 intercept A_0 and the estimates of biases at the medical decision levels along with
256 95% confidence intervals.
- 257 ○ For qualitative assays, provide clinical/analytical performance of the test in the
258 hands of trained operators as sensitivity ($\text{Se}_{\text{Trained}}$) and specificity ($\text{Sp}_{\text{Trained}}$) (or
259 positive percent agreement ($\text{PPA}_{\text{Trained}}$) and negative percent agreement
260 ($\text{NPA}_{\text{Trained}}$)), likelihood ratios, and other applicable measures of performance
261 with 95% two-sided confidence intervals.

262
263 Manufacturers may submit a Pre-Submission to obtain feedback from FDA on planned study
264 designs prior to conducting the study. This may be especially helpful if alternate approaches are
265 planned. For additional information on the Pre-Submission process, please refer to FDA's
266 guidance "Requests for Feedback on Medical Device Submissions: The Pre-Submission Program
267 and Meetings with Food and Drug Administration Staff" available at
268 <https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm311176.pdf>.
269

270
271 Regardless of the option selected, how accuracy is determined will be highly influenced by
272 whether the test is quantitative, qualitative or semi-quantitative.

1) Quantitative Tests

274 A quantitative test is a test that gives numerical results (e.g., concentration of an analyte in a
275 patient sample) which are referenced to a measuring interval and standards. A quantitative test is
276 considered accurate for the purposes of CLIA waiver if the measurement results obtained by
277 untrained operators in CLIA-waived testing settings are comparable to the results of an
278 established quantitative test (i.e., for which results have been traced to references of higher order,
279 usually national or international standards) performed by laboratory professionals. Determination
280 of the comparability of the test results by untrained operators may be made directly (Option 2),
281 or stepwise (Option 1), as long as a conclusion can be made that the candidate test has a
282 negligible likelihood of erroneous results by untrained operators.

283
284 In order to define the accuracy of the candidate test one needs to define the relationship of the value
285 obtained by the test to the true value, and define what contributes to any difference in the values.
286 The following are possible CMs that can be used to determine the accuracy of the candidate test:

- 287
- 288 • Reference Method: The preferred CM is a reference method which has been thoroughly investigated and shown to yield accurate and precise results.
 - 289 • Traceable Calibration Method: If a reference method is not available, an applicant
290 might consider a traceable calibration method that includes calibrators traceable to
291 references of higher order. A reference of higher order can be a certified reference
292 material, a reference measurement procedure, or a network of reference

Contains Nonbinding Recommendations

Draft – Not for Implementation

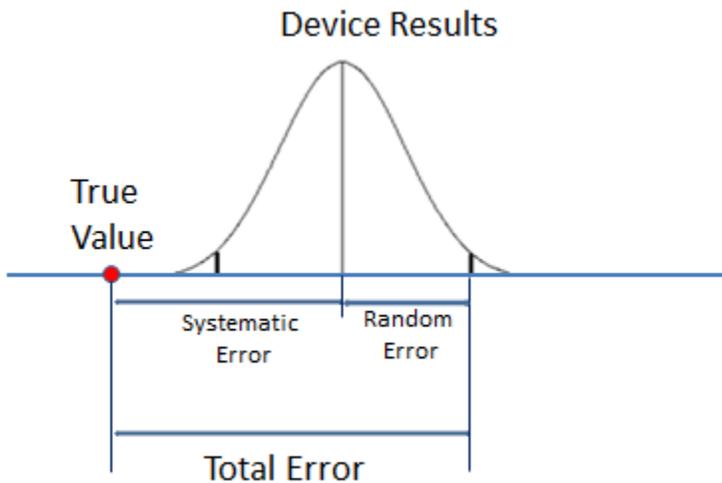
293 laboratories. The results of a traceable calibration method can be related to stated
294 references, usually national or international standards, through an unbroken chain of
295 comparisons in which measurement uncertainties have been documented at every
296 step in the procedure.

297
298 Applicants who choose a traceable calibration method should provide information
299 supporting traceability, including a description of reference materials or methods
300 used to establish the traceability of the calibrators, the values and uncertainties for
301 each of those reference materials or methods, and any other relevant information
302 concerning traceability.

- 303 • Other Well-Documented Method: If neither a reference method nor a traceable
304 calibration method are available, then an applicant may consider another well-
305 documented method. In some situations, the test in the hands of trained operators
306 can be considered as a CM (for example, if the candidate test in the hands of trained
307 operators was evaluated versus a clinical gold standard for the presence/absence of
308 the target condition and there were no other cleared/approved tests on the market).

309
310 In general, the inaccuracy of the test when performed by untrained operators is the deviation (error)
311 of a single measurement from the true value which is obtained by traceable calibration method (or
312 reference method). This includes both the systematic error described by bias and all sources of
313 random errors such as imprecision and random interferences (i.e., patient sample-related random
314 bias). Total error is the largest absolute possible error with 95% confidence and includes all random
315 and systematic errors that can occur during the total testing process from the specimen collection
316 through pre-analytical, analytical and post-analytical phases, see Figure 1.

317
318



319
320 Figure 1. Systematic error, random error and total error.

321
322 All tests have bias and random error, so one needs to define whether the deviation of a single
323 measurement from true value (error) that includes bias and random error is clinically significant.

Contains Nonbinding Recommendations

Draft – Not for Implementation

324 A way to incorporate that into the statistical analysis is by defining zones of allowable total error
325 where 95% of the test values should fall, and zones where no test values should fall, zones of
326 limits of erroneous results. Defining the limits on the allowed error incorporates into the
327 numerical analyses clinical judgment on the benefits and risks posed by the test when used as
328 intended.

329

- 330 • *The Allowable Total Error (ATE) zones.*

331 The ATE zones represent a range of values that can be tolerated without invalidating the
332 medical usefulness of the test. Typically at least 95% of the sample results should fall within
333 the established ATE zones.

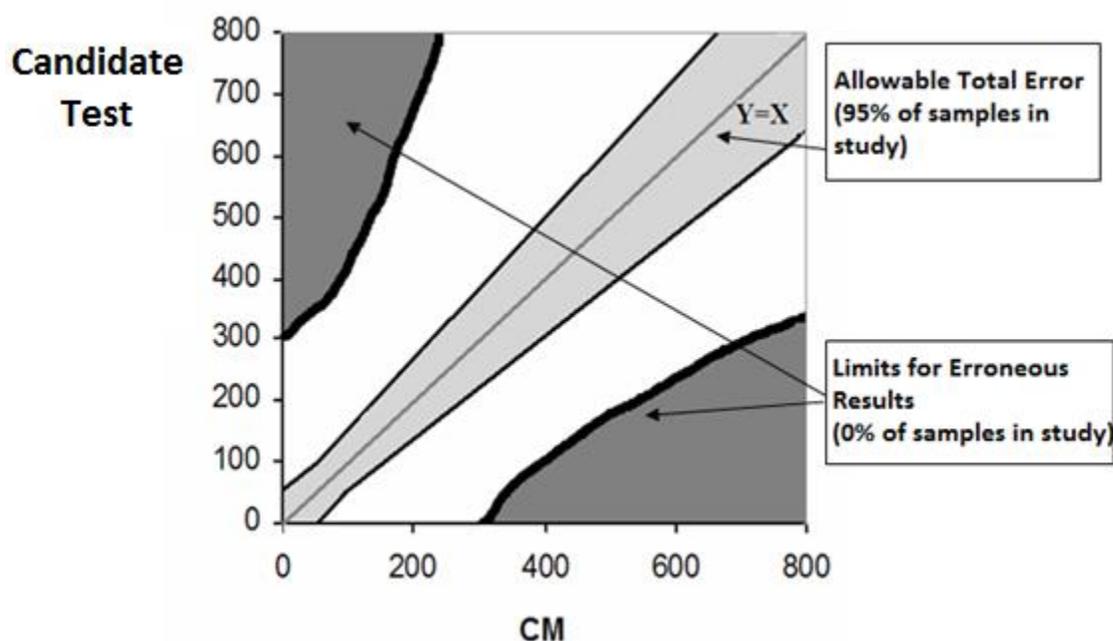
334

- 335 • *The Zones of Limits for Erroneous Results (LER).*

336 While it may be acceptable for a small percentage of results to fall outside the ATE zone,
337 such results should remain within wider limits. Accordingly, applicants should specify
338 Limits of Erroneous Results (LER) zones, within which it is not acceptable for any results to
339 occur as this would pose a risk to patient safety.

340

341 When graphed, such zones give you an error grid as seen in Figure 2.



342

343

344

345 Figure 2. Example of ATE and LER zones.

346

347 Applicants should establish zones of ATE and LER prior to initiating a study.

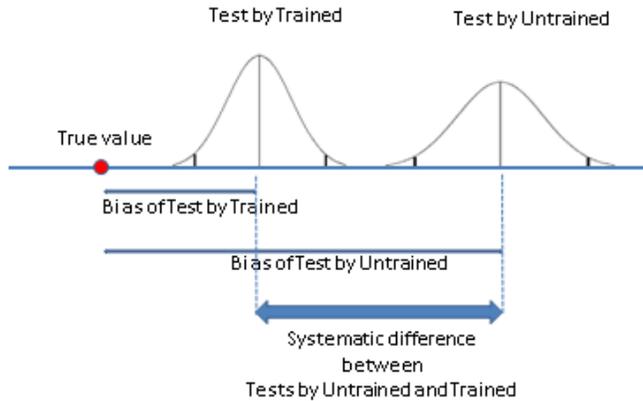
348

349 For Option 1, in the agreement study, each of the subjects should be tested by the candidate test
350 twice: once by untrained operators and once by trained operators. These data are used for
351 estimation of systematic difference between test results by untrained operators versus trained

Contains Nonbinding Recommendations

Draft – Not for Implementation

352 operators at medically important levels (see Figure 3 below). In addition, evaluation of
353 imprecisions of the test when run by untrained and trained operators should be performed for three
354 concentrations (low, medium, and high) and these data are used for comparing test imprecisions in
355 the hands of untrained and trained operators.



356
357 Figure 3. Systematic difference between untrained and trained operators
358

359 For Option 2, possible study designs and analysis plans for quantitative tests should align with
360 the Clinical and Laboratory Standards Institute (CLSI) documents EP21 “Evaluation of Total
361 Analytical Error for Quantitative Medical Laboratory Measurement Procedures.”³EP09⁴ and
362 EP17⁵. The measuring interval of the CM should be at least as wide as the measuring interval of
363 the candidate test and the CM should be run in laboratory settings by trained operators.

364
365 For both options, testing samples across the measuring interval of the candidate test in the study is
366 recommended.

367
368 FDA recommends that applicants interact with FDA on their proposed study design, establishing
369 ATE and LER through the Pre-Submission process prior to initiating a study.

2) Qualitative Tests

370
371
372 A qualitative test is a test that provides two outputs (e.g., positive/negative or yes/no) or multiple
373 nominal⁶ categories. A qualitative test is considered accurate for the purposes of CLIA waiver if
374 the results obtained by untrained operators in CLIA waived settings are comparable to the results
375 of a gold standard (i.e., the best available method for determining whether a target condition is
376 present or absent), or to the results of a quantitative traceable calibration test with an appropriate

³ CLSI. *Evaluation of Total Analytical Error for Quantitative Medical Laboratory Measurement Procedures*. 2nd ed. CLSI guideline EP21. Wayne, PA: Clinical and Laboratory Standards Institute, 2016.

⁴ CLSI. *Measurement Procedure Comparison and Bias Estimation Using Patient Samples; Approved Guideline – Third Edition*. CLSI document EP09-A3, Wayne, PA: Clinical and Laboratory Standards Institute, 2013.

⁵ CLSI. *Evaluation of Detection capability for Clinical Laboratory Measurement procedures; Approved Guideline – Second Edition*. CLSI document EP17-A2, Wayne, PA: Clinical and Laboratory Standards Institute, 2012

⁶ Nominal categories are categories with no intrinsic ordering. For example, an IVD test for genotyping HCV that gives results of multiple categories as 1a, 1b, 2, 3, 4, 5 and 6 is a qualitative test.

Contains Nonbinding Recommendations

Draft – Not for Implementation

377 cutoff, or to the results of a qualitative traceable cutoff test (i.e., when cutoff values have been
378 traced to references of higher order) performed in a laboratory setting by trained operators.
379 Determination of the comparability of the test results by untrained operators may be made
380 directly (Option 2), or stepwise (Option 1), such that a conclusion can be made that the candidate
381 test has a negligible likelihood of erroneous results by untrained operators.

382
383 For qualitative tests, information about bias and random error may not be directly obtainable. In
384 general, the accuracy of a qualitative test is determined by comparison of the test results with one of
385 the comparative methods (CM) described below:

- 386 • **Gold Standard or Qualitative Reference Method:** The preferred CM is a gold standard
387 (a best available method for determining whether the target condition is present or absent
388 (for example, composite reference method, a clinical diagnosis determined by definitive
389 clinical methods), or a qualitative reference method for determination of the presence or
390 absence of the analyte of interest (for examples of qualitative reference methods, consult
391 with OIR)).
- 392 • **Quantitative Reference Method:** a quantitative reference method such as those outlined in
393 the section on quantitative tests, with the appropriate cutoff value for positive and negative
394 results.
- 395 • **Traceable Calibration Method:** A quantitative traceable calibration method with
396 measurement results traceable to higher-order references such as that outlined in the section
397 on quantitative tests with the appropriate cutoff value for positive and negative results.
- 398 • **Traceable Cutoff Method:** A qualitative traceable cutoff method with cutoff
399 determination traceable to higher-order references.
- 400 • **Other Well-Documented Method:** If neither a reference method nor a traceable calibration
401 method are available, then an applicant may consider another well-documented quantitative
402 or qualitative method. For example, this could be a method that was tested by reference
403 specimen panels (e.g., panels of samples prepared by well recognized institutions, such as
404 WHO, CDC, NIST).

405
406 If a gold standard is not available, the performance of a qualitative test may be assessed relative
407 to a quantitative CM with an appropriate cutoff for determination of positive and negative results
408 by the quantitative CM. The quantitative CM should either be a reference method or a traceable
409 calibration method. If such a CM is not available, the applicant may consider a qualitative
410 traceable cutoff method or other well-documented method, such as use of reference specimen
411 panels. We recommend discussing alternatives to quantitative CMs with FDA through the Pre-
412 Submission process prior to initiating such a study.

413
414 In general, accuracy of a qualitative test with two outputs is described by the performance
415 metrics presented below based on the test results and gold standard results for N representative
416 subjects from the target population:

417
418

419
420
421

Table 1. Qualitative Test vs. Gold Standard

		Gold Standard		Total
		Target Condition present	Target Condition absent	
Test	Pos	A	B	A+B
	Neg	C	D	C+D
Total		N ₁ =A+C	N ₀ =B+D	N

422
423
424
425
426
427
428
429
430
431
432

- Clinical Sensitivity (Se) = A/(A+C)
- Clinical Specificity (Sp) = D/(B+D)
- Positive Likelihood Ratio (PLR) = (A/N₁)/(B/N₀)
- Negative Likelihood Ratio (NLR) = (C/N₁)/(D/N₀)
- Prevalence (π)⁷ = N₁/N
- Positive Predictive Value (PPV) = A/(A+B)
- Negative Predictive Value (NPV) = D/(C+D)
- Percent of subjects with Test positive results = (A+B)/N
- Percent of subjects with Test negative results = (C+D)/N

433
434
435
436

If the gold standard is not available, test performance metrics are positive and negative percent agreements described below based on the test and comparative method results for N representative subjects from the target population:

437 **Table 2. Qualitative Test vs. Quantitative Test**

		CM		Total
		Positive	Negative	
Test	Pos	A	B	A+B
	Neg	C	D	C+D
Total		N ₁ =A+C	N ₀ =B+D	N

438
439
440

- Positive Percent Agreement (PPA) = A/(A+C)
- Negative Percent Agreement (NPA) = D/(B+D)

441
442

The formulas for sensitivity, specificity, and positive and negative percent agreements appear to be the same; however, it is critical to recognize that these values are not interchangeable⁸.

⁷ If the prevalence in the study was different from the prevalence in the target population, then PPV and NPV for this prevalence can be calculated using the following relationship between predictive values, likelihood ratios and prevalence as: $\frac{PPV}{1-PPV} = PLR \cdot \frac{\pi}{1-\pi}$ and $\frac{1-NPV}{NPV} = NLR \cdot \frac{\pi}{1-\pi}$.

⁸ CLSI. *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline – Second Edition*. CLSI document EP12-A2, Wayne, PA: Clinical and Laboratory Standards Institute, 2008.

Contains Nonbinding Recommendations

Draft – Not for Implementation

443 Applicants should establish criteria for acceptable levels of performance prior to beginning a
 444 study regardless of the CM selected. Criteria should be based on a benefit-risk analysis taking
 445 into consideration benefits for patients (e.g., availability of results at the time of the doctor’s
 446 office visit) and risks (e.g., incorrect results when performed by untrained operators). FDA
 447 generally recommends that criteria for sensitivity (or PPA between test and CM) and specificity
 448 (or NPA between test and CM) should be 95% or greater; and that the lower confidence bound
 449 for the two-sided 95% score confidence interval should be at least 89%. In some cases, tighter
 450 performance criteria may be needed to reasonably assure that the candidate test is “accurate” and
 451 in other cases lower performance criteria may be clinically acceptable with sufficient benefit-risk
 452 justification.
 453

454
 455 For Option 1, in the agreement study, each of the subjects should be tested by the candidate test
 456 twice: once by untrained operators and once by trained operators. These data are used for
 457 calculation of PPA and NPA between test performed by untrained operators and test performed by
 458 trained operators:
 459

Table 3. Test by Untrained Operators vs. Test by Trained Operators

		Test by Trained		Total
		Positive	Negative	
Test by Untrained	Pos	A	B	A+B
	Neg	C	D	C+D
Total		$N_1=A+C$	$N_0=C+D$	N

461
 462 Estimate the $PPA_{Untrained|Trained}$ and the $NPA_{Untrained|Trained}$ as
 463 $PPA_{Untrained|Trained}=A/(A+C)$

464 $NPA_{Untrained|Trained}=D/(B+D)$
 465

466 If a numeric signal of the test is available, then you can calculate zones around the cutoff for
 467 which samples may have discordant results if run multiple times by trained operators. Using
 468 these zones, we recommend that you present results of the agreement study in a detailed table of
 469 agreement:
 470

Table 4. Test by Untrained Operators vs. Numeric Signal of Test by Trained Operators

		Test by Trained				Total
		Positive		Negative		
		High and Moderate Positive	Low Positive (close to cutoff)	High Negative (close to cutoff)	Low and Moderate Negative	
Test by Untrained	Pos					
	Neg					
Total		N_1		N_0		N

Contains Nonbinding Recommendations

Draft – Not for Implementation

473 Taking into consideration acceptable levels of sensitivity (or PPA between test and CM) and
474 specificity (or NPA between test and CM) and levels of test accuracy in the pre-market
475 submission, you should evaluate whether the levels of agreement between test results in the
476 hands of untrained and trained operators in the agreement study are acceptable.

477
478 Usually, an agreement study does not have enough subjects close to the cutoff that test performance
479 with untrained and trained operators for the subjects close to the cutoff can be evaluated. Therefore,
480 in addition, comparison of imprecisions of the test when run by untrained and trained operators
481 should be performed for concentrations C_5 and C_{95} ⁹ (or limit of detection) close to the cutoff.

482
483 For Option 2, the study should include subjects from the intended use population and these subjects
484 should be tested by untrained operators and results should be compared with results from the CM
485 performed by trained operators. In addition, samples near the cutoff (i.e., C_5 and C_{95} of the CM if
486 the CM is a quantitative test or C_5 and C_{95} of the candidate test if the CM is a gold standard for the
487 target condition) should be tested by untrained operators. For analysis plans for qualitative tests, see
488 the Clinical and Laboratory Standards Institute (CLSI) document EP12.¹⁰

489
490 For Options 1 and 2, we recommend discussing study sample with FDA through the Pre-
491 Submission process prior to initiating a study.

492

493 **3) Semi-Quantitative Tests**

494 For the purpose of this guidance, a semi-quantitative test is a test with a few ordinal categories
495 (e.g., negative, trace, +, ++, +++) in which the order of categories together with the definitions of
496 these categories contain information used during the interpretation of the test results. A semi-
497 quantitative test is considered accurate for the purposes of CLIA waiver if the results obtained by
498 untrained operators in CLIA-waived testing settings are comparable to the results obtained by
499 trained operators performing a traceable calibration test (see “Quantitative Tests” section above
500 for more information), with appropriate cutoffs for defining ordinal categories

501

502 An example of a semi-quantitative test with four ordinal categories (negative, trace, +, ++) is
503 presented in Figure 4.

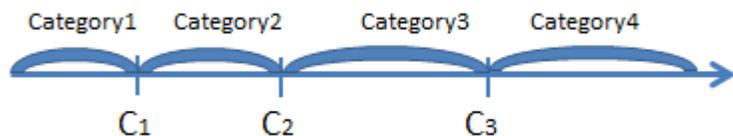
504

⁹ CLSI. *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline – Second Edition*. CLSI document EP12-A2, Wayne, PA: Clinical and Laboratory Standards Institute, 2008.

¹⁰ CLSI. *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline – Second Edition*. CLSI document EP12-A2, Wayne, PA: Clinical and Laboratory Standards Institute, 2008.

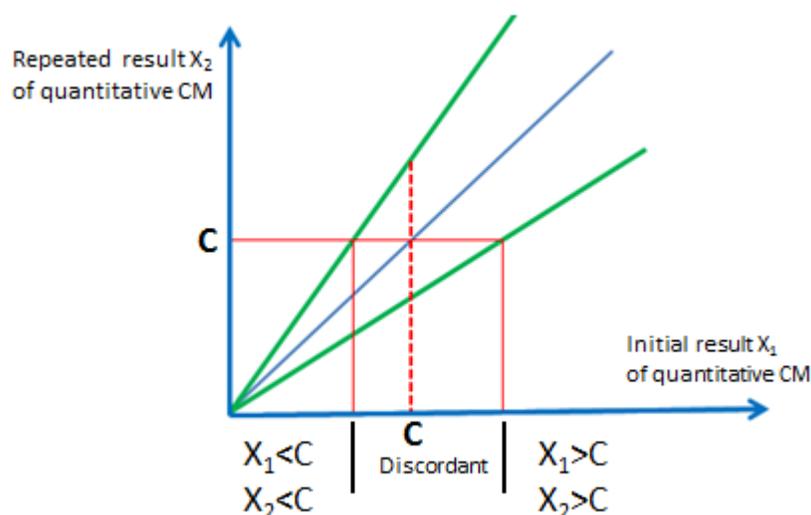
Contains Nonbinding Recommendations

Draft – Not for Implementation



505
506 Figure 4. Semi-quantitative test with 4 categories based on three (3) cutoffs for numeric values.

507
508 The performance of a semi-quantitative test may be assessed relative to a quantitative
509 comparative method (CM). The quantitative CM should either be a reference method or a
510 traceable calibration method. Any quantitative CM chosen will have some random error
511 associated with it such that two measurements of the same sample may differ slightly. Thus, a
512 sample close to a cutoff C of the candidate semi-quantitative test may have an initial result above
513 the cutoff C and subsequent result below the cutoff C . As samples close to cutoff C may have
514 discordant results when tested multiple times by the quantitative CM, one can anticipate that
515 these samples may also have discordant categories when tested by the candidate semi-
516 quantitative test. A zone¹¹ around each cutoff can be established based on the allowable
517 imprecision of the quantitative CM (%CV or SD) such that 95% of the differences between two
518 repeated measurements of the quantitative CM, X_2 and X_1 , are inside this zone (see Figure 5).



519
520
521 Figure 5. Two repeated measurements from quantitative test.

522
523 Applicants should establish near-cutoff ranges around each cutoff using the 95% zone as
524 described above. For samples with measurements by the CM within one of these ranges, it is
525 expected that candidate semi-quantitative test results may fall into either of the categories
526 bordering the cutoff. For CM test results outside of the near-cutoff ranges only one

¹¹ These boundaries are symmetrical around the diagonal when the distance is calculated perpendicular to the diagonal.

Contains Nonbinding Recommendations

Draft – Not for Implementation

527 corresponding bin of the semi-quantitative test is allowed. To summarize, the ATE zones for
 528 semi-quantitative candidate tests are the allowable bins for the different ranges of the
 529 quantitative CM values. For CM results in the near-cutoff range, both semi-quantitative
 530 candidate test bins are allowable and only one corresponding bin of the semi-quantitative
 531 candidate test is allowable for the CM values outside of these near-cutoff ranges. The LER zones
 532 are defined as the bins of the semi-quantitative candidate test that differ from the allowable bins
 533 by more than one bin.

534 This analysis with ATE and LER zones may be presented as a table of agreement between semi-
 535 quantitative and quantitative tests as shown in Table 5 where the light grey and dark grey shaded
 536 cells represent the ATE and LER zones correspondingly.

537

538 **Table 5. Agreement of Semi-Quantitative Test and Quantitative CM**

539

		Quantitative CM							
			Near cutoff C ₁		Near cutoff C ₂		Near cutoff C ₃		Total
Expected Test Categories		Only Cat.1	Cat.1 or Cat.2	Only Cat.2	Cat.2 or Cat.3	Only Cat.3	Cat.3 or Cat.4	Only Cat.4	
Test	Cat.1	A ₁	B ₁	C ₁	D ₁	E ₁	F ₁	G ₁	
	Cat.2	A ₂	B ₂	C ₂	D ₂	E ₂	F ₂	G ₂	
	Cat.3	A ₃	B ₃	C ₃	D ₃	E ₃	F ₃	G ₃	
	Cat.4	A ₄	B ₄	C ₄	D ₄	E ₄	F ₄	G ₄	
Total		A	B	C	D	E	F	G	N
Observed percent in ATE		A ₁ /A	(B ₁ +B ₂)/B	C ₂ /C	(D ₂ +D ₃)/D	E ₃ /E	(F ₃ +F ₄)/F	G ₄ /G	
		Total percent in ATE= $(A_1+B_1+B_2+C_2+D_2+D_3+E_3+F_3+F_4+G_4)/N$ Total percent in LER= $(A_3+A_4+B_4+C_4+E_1+F_1+G_1+G_2)/N$							

540

541 For Option 1, in the agreement study, each of the subjects should be tested by the candidate test
 542 twice: once by untrained operators and once by trained operators. These data are used for
 543 calculation of percent agreements between test performed by untrained operators and test performed
 544 by trained operators. For example, for N samples from the intended use population in the
 545 agreement study, the performance of a semi-quantitative candidate test with four categories may
 546 be presented as described in Table 6. The grey shaded cells in Table 6 represent samples for
 547 which there was agreement between the test results by untrained and trained operators.

548

549 **Table 6. Comparison of Two Semi-Quantitative Tests**

550

		Test by Trained				
		Category1	Category2	Category3	Category4	Total

Contains Nonbinding Recommendations

Draft – Not for Implementation

Test by Untrained	Category1	A ₁	B ₁	C ₁	D ₁	
	Category2	A ₂	B ₂	C ₂	D ₂	
	Category3	A ₃	B ₃	C ₃	D ₃	
	Category4	A ₄	B ₄	C ₄	D ₄	
Total		A ₁ +A ₂ +A ₃ + A ₄	B ₁ +B ₂ +B ₃ + B ₄	C ₁ +C ₂ +C ₃ + C ₄	D ₁ +D ₂ +D ₃ + D ₄	N

551
552 Performance of the semi-quantitative test is described by the percent agreements (PA) of the
553 categories of the test in the hands of untrained operators with categories of the test in the hands
554 of trained operators as:

- 555
- 556 • $PA_{\text{Category1,Untrained|Category1,Trained}} = A_1 / (A_1 + A_2 + A_3 + A_4)$
 - 557 • $PA_{\text{Category2,Untrained|Category1,Trained}} = A_2 / (A_1 + A_2 + A_3 + A_4)$
 - 558 • $PA_{\text{Category3,Untrained|Category1,Trained}} = A_3 / (A_1 + A_2 + A_3 + A_4)$
 - 559 • $PA_{\text{Category1,Untrained|Category2,Trained}} = B_1 / (B_1 + B_2 + B_3 + B_4)$ and so on.
- 560

561 For Option 2, the study includes subjects from the intended use population and these subjects should
562 be tested by untrained operators and results should be compared with results from the CM
563 performed by trained operators. The CM should be either a reference method or a traceable
564 calibration method. ATE and LER should be established and analysis should be performed as
565 described above. We recommend discussing selection of the CM with FDA through the Pre-
566 Submission process prior to initiating such a study.

567

568 For Options 1 and 2, in addition, samples near each of the clinically important cutoffs should be
569 tested. The number of samples needed will depend on the number of categories and the importance
570 of each category. For Option 1, the imprecisions of the candidate test when performed by
571 untrained and trained operators should be compared by testing C₅ and C₉₅ samples for each
572 clinically important cutoff. For Option 2, samples close to the cutoffs (i.e., near C₅ and C₉₅ of
573 quantitative CM) should be tested by untrained operators for clinically important cutoffs.¹²

574

575 It should be noted that when a semi-quantitative test is compared to the same or another semi-
576 quantitative test in a study and the level of agreement is not high, this study design does not
577 provide enough information to evaluate whether the levels of disagreement are acceptable. This
578 is because one cannot determine whether the samples with discordant results were close to the
579 cutoffs (i.e., where some level of disagreement would be expected) or far from the cutoffs (i.e.,
580 where a very high level of agreement is anticipated). Accordingly, Option 1 is recommended
581 only when a quantitative CM is not available and a high level of agreement of the test results
582 between untrained and trained operators is expected.

583

584 For Options 1 and 2, we recommend discussing study size with FDA through the Pre-Submission
585 process prior to initiating a study.

¹² CLSI. *User Protocol for Evaluation of Qualitative Test Performance; Approved Guideline – Second Edition*.
CLSI document EP12-A2, Wayne, PA: Clinical and Laboratory Standards Institute, 2008.

4) General Study Design Considerations

For all study designs, FDA recommends that applicants evaluate test performance in settings designed to replicate, as closely as possible, the actual CLIA waived settings and patients/samples. Therefore, study designs should include the following:

- CLIA Waiver testing sites that are representative of the intended use population.
- CLIA Waiver intended operators (untrained operators). We encourage you to enroll operators with the least amount of training that might be encountered at the types of sites for which this device is intended.
- Intended sample type and matrix.
- Testing over time, as in the typical intended use setting.

a) Testing sites

You should conduct the study to support CLIA waiver at a minimum of three sites that are representative of both the intended use patient population and the intended operators in CLIA waived settings. Generally, the sites should include different demographic locations (e.g., outpatient clinic, physician's office), since patient populations and intended operators typically vary between different demographic locations. In your CLIA waiver application, you should present a brief description of each site, including its name, address, and the date the study was performed. If there were sites that were included at the beginning, but then did not complete the study, you should provide a brief explanation for why those sites did not complete the study.

b) Study participants

1. Operators

a) CLIA Waiver intended operators (untrained operators)

The study should include 1-3 untrained operators at each site and at least nine (9) untrained operators across all sites. You should ensure that the untrained operator study participants enrolled represent anticipated operators of the device you propose for CLIA waiver. We recommend that you record and tabulate the education (including experience and training) and the occupation of each operator to demonstrate that these participants meet the definition of intended operators and include this in your CLIA waiver application. In addition, for each study site, we ask you to report the same information on other operators that were available at the testing site but that were not chosen to participate.

Contains Nonbinding Recommendations

Draft – Not for Implementation

623 b) **Trained operators**

624

625 The study should include at least one laboratory professional (trained operator) at each site.
626 These laboratory professionals should meet the qualifications to perform moderate complexity
627 testing and have previous training in performing the candidate test.

628 c) **Instructions for use**

629

630 You should provide the untrained operators who participate in the study with only the Quick
631 Reference Instructions (see Section VI.C. Educational Information of the guidance entitled
632 “Recommendations: Clinical Laboratory Improvement Amendments of 1988 (CLIA) Waiver
633 Applications for Manufacturers of In Vitro Diagnostic Devices”
634 (<https://www.fda.gov/RegulatoryInformation/Guidances/ucm079632.htm>). The untrained
635 operators should receive no additional instructions (e.g., written or verbal training, coaching, or
636 prompting). Likewise, untrained operators should have no opportunity to discuss the test with
637 other participants or otherwise coach or observe each other. Untrained operators may call an 800
638 number help-line if such a service is to be provided for the device when it is marketed. You
639 should include, in your waiver application, the instructions you provided to untrained operators
640 participating in the study.

641

642 d) **Universal precautions**

643

644 You should comply with the Federal Food, Drug, and Cosmetic Act (FD&C Act) and its
645 implementing regulations and should ensure your study complies with all other pertinent laws
646 and regulations, including Occupational Health and Safety Administration (OSHA) regulations
647 pertaining to biological hazards (“universal precautions”), 29 CFR 1910.1030.

648

649 e) **Operator questionnaire**

650

651 You should develop an operator questionnaire to be filled out by all untrained operators
652 participating in the study. This questionnaire should be designed to help assess whether the
653 untrained operators understood how to use the device correctly. It is important that the
654 questionnaire be given to test untrained operators *after* the completion of the clinical study, so
655 the questions do not bias the untrained operators during the study. Some questions may ask
656 untrained operators to indicate agreement on a 1-5 scale (1=strongly disagree; 5=strongly agree).
657 The following are examples:

658

- 659 • The instructions were easy to follow.
- 660 • It was easy to apply the sample correctly.
- 661 • It was easy to see and understand the test results (e.g., appearance of the line, change
662 of color).
- 663 • The control line was always distinct and easy to read.

Contains Nonbinding Recommendations

Draft – Not for Implementation

- 664 • The instructions clearly explain what to do if a test result does not appear or is
665 invalid.
- 666 • I needed help from someone the first time I ran the test.
667

668 We recommend that, as part of the questionnaire, you show various possible test results and
669 control results that are positive, negative, and invalid and ask the untrained operator to read these
670 results. You may wish to present these questions as true/false or multiple choice questions.

671 You should also strongly encourage general comments by the untrained operators. We
672 recommend that you include your survey questions and results with your CLIA application.

673 c) **Subjects (Patients) and Samples**

674

675 In order to prevent biases, samples should be collected from consecutive patients over one
676 month. Depending on the specific clinical site, the prevalence of the disease, or other factors, it
677 may be appropriate to limit consecutive enrollment to two (2) weeks.
678

679 You should ensure that subjects from whom you will obtain specimens for the clinical study
680 meet inclusion and exclusion criteria corresponding to the intended use population of the test.
681 Once a subject has been determined to meet appropriate inclusion criteria, he/she should be
682 informed of the study and invited to participate.
683

684 You should follow applicable laws and regulations for human subject protection, including
685 patient privacy and informed consent. See section 520(g) of the FD&C Act; 21 CFR parts
686 50,56, and 812; and the Health Insurance Portability and Accountability Act (HIPAA) [P.L. 104-
687 191]; 45 CFR Part 46.
688

689 Samples should adequately represent all possible values of the test. If possible, applicants should
690 strive to achieve this at each site as well as across all sites. For quantitative and semi-
691 quantitative candidate tests, samples should span the measuring intervals of the device and study
692 data should include a few samples around Medical Decision Levels (MDLs). For qualitative test,
693 samples in the study should include samples near the cutoffs.
694

695 Each sample should be split in two parts. One part should be tested by the untrained operator
696 using the test and the other part should be tested by a trained operator using the test for Option 1
697 and by CM for Option 2. If the sample cannot be split into parts, then a second sample from the
698 same patient should be collected within a suitable time interval. We recommend consultation
699 with FDA through the Pre-Submission process if the order in which the samples are collected
700 impacts the results of testing. Untrained and trained operators should be blinded to test results
701 from other operators and other samples.
702

703 We recommend using actual patient specimens to best assess a device in the hands of untrained
704 operators. However, in some situations, when samples from some categories are rare, it may be
705 appropriate to supplement prospective patient samples with archived samples. If archived
706 patient samples are not available, it may be appropriate to supplement patient samples with

Contains Nonbinding Recommendations

Draft – Not for Implementation

707 contrived samples, such as individual spiked or diluted patient samples. Spiked, diluted, or
708 otherwise contrived samples used in the study should be individual samples (i.e., they should not
709 be aliquots from a single pool). Any archived or contrived sample matrix should be the same as
710 that of the intended use patient samples. Applicants should describe the origin of such samples
711 and how they were prepared. For qualitative and semi-quantitative tests, archived and contrived
712 samples should include samples near the cutoffs. Use of archived or contrived samples should be
713 appropriately justified. In general, archived or contrived samples should not comprise greater
714 than one third of the total study samples; however, there may be some situations in which more
715 or less would be appropriate when an adequate justification is provided. The patient and
716 contrived samples should be as equally distributed among the untrained operators as possible.
717 FDA encourages applicants to discuss planned use of archived or contrived samples with OIR in
718 advance.
719

720 d) **Financial disclosure**

721 If clinical investigators are involved in the clinical study, you should include a Financial
722 Disclosure Statement with your waiver application. For information on financial disclosure
723 statements, we recommend you consult, “Guidance for Industry: Financial Disclosure by Clinical
724 Investigators,” available at
725 <https://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM341008.pdf> and 21
726 CFR Part 54, Financial Disclosure by Clinical Investigators.

727 e) **Clinical study reports**

728
729 You should report results of the clinical study intended to support CLIA waiver by each intended
730 site and overall, if appropriate. Reports should include the following:
731

- 732 • Protocol description.
- 733 • Number of subjects (i.e., patients) studied.
- 734 • Procedures for subject inclusion and exclusion.
- 735 • Description of the subject population.
- 736 • Description of how specimens were collected and stored.
- 737 • Masking techniques.
- 738 • Discontinuations.
- 739 • Complaints, device failures, and replacements.
- 740 • Any invalid results and how these were handled.
- 741 • Information about QC procedures that were performed.
- 742 • Pertinent tabulations.
- 743 • Annotated line listings of results (including electronic versions).
- 744 • Clear descriptions and presentations of the statistical analyses.
- 745 • An explanation for data that are incomplete or missing (Note: You should not
746 remove “outliers”).

Contains Nonbinding Recommendations

Draft – Not for Implementation

747
748
749
750
751
752
753
754
755
756
757
758
759
760
761

You should also report the following for each untrained and trained operator:

- Total number of performed candidate tests.
- Number of initial invalid results.
- Number of retested results.
- Number of final invalid results.

You should calculate and report the percentage of initial and final (if applicable) invalid results with a 95% two-sided confidence interval and then exclude invalid results from calculations of the test performance characteristics. Please provide a rationale as to why the observed percentage of invalid results is clinically acceptable.

DRAFT