Consult Review Memo

FDA/CDRH/OIR

From: Noel Gerald, Ph.D., Division of Microbiology Devices, OIR, CDRH

To: Christopher Leptak, Ph.D., Office of New Drugs, CDER

Date: Nov 7, 2018

Consult: DDTBMQ000044
FDA, CDER, Drug Development Tool Program, Biomarker Qualification Review Package

Product: *Plasmodium* 18S rRNA/rDNA biomarker
Applicant: Sean C. Murphy, MD/Ph.D.
University of Washington, Department of Laboratory Medicine
750 Republican Street, E633
Seattle, Washington 98109

Through: Uwe Scherf, M.Sc., Ph.D., Director, Division of Microbiology Devices, OIR, CDRH

Background:
CDER requested a CDRH consult for a Full Qualification package submitted to the CDER, Drug Development Tool (DDT) Biomarker Qualification (BQ) program. The package was submitted by Sean Murphy, MD/Ph.D., U. of Washington, to qualify *Plasmodium* 18S rRNA/rDNA as a biomarker for use in the controlled human malaria infection studies (CHMI) that may be included in regulatory submissions for antimalarial drug (CDER) or vaccine (CBER) products. The BQ review team includes members from CDER, CBER, and CDRH.

CHMI
In CHMI, small numbers of healthy, malaria-naïve volunteers are inoculated with an infectious laboratory strain of *P. falciparum* under controlled conditions. The parasites (blood stages or liver-infecting stages) can be inoculated using needle injection or mosquito bite. The subjects are typically divided among control and treatment arms to assess the ability of an experimental drug or vaccine product to prevent or treat blood stage *P. f.* infection. Historically, CHMI subjects receive rescue antimalarial drug therapy as a safety measure per predefined criteria such as the recognition of malaria symptoms, the identification of blood stage parasites by thick blood smear (TBS) microscopy, or at the end of the observation period, regardless of evidence of infection. Depending on the study design and objective, subjects in the experimental treatment group that require rescue therapy may be considered “failures” with respect to the experimental treatment.

Asexual blood stage parasite infection
The asexual blood stages of *Plasmodium* parasites are responsible for the symptoms of malaria, and the peripheral blood levels of these parasites fluctuate as the parasite burden increases

during infection. Parasite peripheral blood levels fluctuate due to the combined effects of synchronized parasite growth, periodic sequestration of parasites from the blood into the vasculature as they mature during the asexual life cycle, and the clearance of parasites from the blood by the host immune system.

Malaria symptom onset depends in part on patient age, immune status, prior exposure history, and parasite-specific factors. Malaria-naïve individuals can begin to experience symptoms when parasites are detectable by TBS microscopy in the blood, and importantly, they can begin to experience symptoms before they are positive by TBS microscopy. For these reasons, CHMI subjects have often been housed in hotels to allow frequent safety monitoring for symptoms and TBS status during the study.

Purpose of submission:
The purpose of the current package is to qualify *Plasmodium* 18S rRNA/rDNA, detected by a nucleic acid amplification test (NAAT), as a biomarker to replace TBS microscopy in CHMI submissions to CDER and CBER.[1]

Context of Use (COU) as modified during internal BQ review team discussion
Biomarker: *P. falciparum* 18S rRNA/rDNA, A type
COU: A monitoring biomarker to inform initiation of rescue treatment with an anti-malarial drug following controlled human malaria infection (CHMI) with *P. falciparum* sporozoites in healthy subjects from non-endemic areas enrolled in clinical studies for vaccine and drug development against *P. falciparum*.

Summary Observations and Conclusions:

- Nucleic acid testing is a fundamentally different technology than microscopy, and so it raises new questions for CHMI that are not encountered with TBS microscopy. These include new questions about the clinical relevance of some NAAT results in CHMI.

- The applicant proposes analytical performance claims for quantitative biomarker NAATs that are difficult to confirm, and in some instances the CDRH review did not agree with the applicant's results presentation or conclusions at the lowest biomarker levels. However, CDER/CBER may not consider such disagreements at low biomarker levels to be clinically relevant.

- In the absence of consensus actionable decision points that define what parasite nucleic acid concentrations are considered clinically relevant, there is little basis for concluding that the described analytical performance is not sufficient for the COU. The review comments below should not preclude the qualification of the parasite 18S rRNA/rDNA biomarker.

- The information provided supports the claim that NAAT detection of parasite 18S rRNA/rDNA as a biomarker in the blood can be a more sensitive indicator of blood stage parasite infection than the detection of parasites by TBS microscopy. This has the

[1] Note: Biomarker qualification is not strictly required for the use of NAATs in regulatory submissions to CDER and CBER (e.g., IND, NDA, BLA, etc.). Sponsors of regulatory submissions to CDER/CBER currently provide information about relevant NAATs and other testing for review on a case-by-case basis.

potential to identify blood stage infections before CHMI subjects become positive by TBS microscopy or begin to experience symptoms of malaria.

Consult Review Comments:
The following review comments are provided specifically for the biomarker qualification program. These comments are not intended to represent CDRH recommendations or requirements for premarket submissions to CDRH.

Review concepts

1. The proposed biomarker is parasite rRNA/rDNA

   It is important to note that the proposal is to qualify parasite rRNA/rDNA as a biomarker, and that NAATs are proposed to detect and quantitate the biomarker. While the proposal frequently describes the NAATs as reporting parasites/mL, in most cases this is not appropriate. The *Plasmodium* parasite itself should not be considered the biomarker. The quantitative NAATs use armored RNA calibration curves to produce the results, and so the appropriate units for assessing NAAT quantitative performance are biomarker copies/mL. Despite the applicant's tendency to describe NAAT performance as parasites/mL, it would be incorrect to assume that rRNA/rDNA copies are only associated with live, intact parasites in discrete ordinal steps (e.g., 1, 10, 100 parasites/mL). Thousands of rRNA/rDNA copies are predicted per parasite.

   There must be a distinction between biomarker copies/mL and predicted parasites/mL, as this will affect the descriptions of both analytical and clinical study results.

2. Proposal claims

   A key assertion of the proposal is that NAAT detection of parasite 18S rRNA/rDNA as a biomarker in the blood is a more sensitive indicator of blood stage parasite infection than the detection of parasites by TBS microscopy. This is supported by published literature which indicates that NAAT detection of parasite nucleic acids can be 100-fold to 1000-fold more sensitive than the detection of whole parasites by TBS microscopy. The applicant provided a summary of several published studies in which NAATs detected parasite nucleic acids in blood samples days earlier than parasites were identified by TBS microscopy. Having the ability to identify blood stage infections before subjects become positive by TBS microscopy and before the onset of malaria symptoms would improve the safety of CHMI for the subjects.

   However, nucleic acid testing is a fundamentally different technology than microscopy, and the differences can raise new questions of safety and effectiveness for patient testing during CHMI. In this regard, the submission raises new questions and proposes new claims that exceed the scope of the current use of TBS microscopy in CHMI.

   Novel aspects of the proposal that require consideration include using quantitative NAATs to measure *Plasmodium* 18S rRNA/rDNA biomarker concentrations in blood, using biomarker concentration values to extrapolate blood parasite densities at levels that are undetectable by TBS microscopy, and defining threshold levels of parasite

biomarker concentrations as clinically actionable or clinically insignificant in CHMI studies. These aspects are discussed in more detail in comments 3 and 4 below.

3. Detecting biomarker and estimating blood parasite densities below the LoD of TBS microscopy

   The applicant proposes to use NAAT reported rRNA/rDNA biomarker blood concentrations to extrapolate blood parasite densities in the range of 20 parasites/mL.[2] This raises new questions that were not previously considered with the use of TBS microscopy in CHMI, as these values are two to three logs below values reported for the limit of detection (LoD) of TBS microscopy.

   The TBS microscopy LoD will vary by method and the experience of the technician, and is commonly cited in the literature as 50,000 – 100,000 parasites/mL for expert microscopists.[3] Even higher parasite density levels would likely be required for quantitation by TBS microscopy with high confidence. The variability of TBS microscopy quantitation is expected to be greatest at the lower parasite densities detectable by this method, due to factors such as parasite loss during sample processing, sampling effects, and microscopist experience[4].

   As a reminder, quantitation is a separate claim from detection. It is possible for a test to detect a target in a sample qualitatively (positive/negative), without being able to quantitate that target with sufficient accuracy for a stated use. The use of biomarker NAATs in CHMI will raise new questions for both qualitative and quantitative considerations.

   a. Qualitative result considerations

      TBS microscopy-negative samples would typically not be considered actionable in CHMI in the absence of clinical symptoms. However, NAATs will detect parasite nucleic acids in some of these microscopy-negative samples, raising new questions on how to manage these subjects in the study. A subject with a NAAT-positive and TBS-negative result may be managed differently depending on whether they are in the control arm or treatment arm of a CHMI study. For example, when ensuring the safety of control subjects challenged with *P. falciparum* sporozoites in CHMI, CDER/CBER may consider <u>any</u> amount of parasite nucleic acids detected in the blood after a defined date to be indicative of a blood stage infection requiring rescue treatment.

      Due to the biology of the parasite, some fluctuations between positive and negative NAAT results should be expected. In the clinical line data provided,

[2] See separate discussion of sponsor proposed conversion factors in comment 19.
[3] See Appendix I for LoD definition. While the submission reports TBS values as low as 1,700 parasites/mL in the clinical study line data, these nominal values are likely below the levels that can be successfully detected by TBS more than 95% of the time.
[4] Few published studies directly address the variability of parasite quantitation by TBS microscopy; however, up to a two-fold difference in parasite density TBS counts has been reported in samples at 300,000 parasites/mL (Bowers, et al, *Malaria J.,* 2009). Therefore, the 1,700 parasites/mL TBS densities reported in the current submission are also below the levels that can be quantitated with very high confidence using this method.

there were several examples of subjects for whom NAAT results fluctuated positive-negative-positive over three days of sequential testing (Table 1).

Table 1. Clinical Study Subjects with Sequential Pos-Neg-Pos NAAT Results

Clinical Study	Subject(s)
MC-003	5, 6, 7, 16, and 30
CVAC-PYR	050D, 005E, and 19V
MC-001	30-06-2

b. Quantitative results considerations

i. Quantitative claims increase the complexity of the discussion by introducing new questions such as:

- How well does the NAAT quantitate biomarker copies in clinical samples?
- How accurately can NAAT biomarker quantitative results estimate blood parasite densities near actionable decision points?

Although the submission frequently implies that these are the same question, this approach would not be appropriate. As noted, there should be thousands of rRNA/rDNA copies per parasite, and free nucleic acids from dead/lysed parasites can contribute to the count of biomarker copies in the sample. The questions of biomarker quantitation (which includes dead parasites and parasite fractions) should be considered separately from estimates of parasite density.

Therefore, a central question for CDER/CBER regarding the use of quantitative NAATs in CHMI is

- What are the clinically relevant levels of parasite nucleic acids (or changes in parasite nucleic acid levels) that define the actionable decision points in CHMI?

As discussed in comment 3.b.iv below, there does not appear to be a clear consensus for what might be considered the clinically relevant levels of parasite nucleic acids in CHMI.

ii. A comprehensive assessment of quantitative NAAT performance would first require actionable decision points to be defined

To evaluate the accuracy of nucleic acid quantification by a NAAT, it would be necessary to first define clinically actionable decision points for a specific use. The actionable decision points are needed to designate the maximum error allowed for the test, as the acceptable error will depend on how the results will be used. It would not be appropriate to use analytical performance data as the sole (or majority) support to define a clinically actionable decision point.

As there are currently no well-defined actionable decision points for quantified *Plasmodium* nucleic acid levels, there is little basis for determining whether an analytical measurement error is clinically relevant. Therefore, the review discussions of analytical performance below are largely descriptive.

iii. Challenges for verifying sub-microscopic parasite densities

It is difficult to establish how accurately NAAT quantitation of parasite nucleic acids may estimate blood parasite densities at levels that are below the LoD of TBS microscopy (sub-microscopic). There are both technical and biological limitations to these analytical studies, and so experimental artifact will become an increasing concern for studies attempting to validate very low sub-microscopic claims.

In general, NAAT performance claims should be based on what can be demonstrated experimentally rather than theoretically, and the justifications for acceptable performance should be tied to clinical relevance. As described below, it does not appear that all the applicant's quantitative claims can be demonstrated experimentally by the analytical study results, particularly at the lowest levels claimed.

However, CDER/CBER may ultimately conclude that quantitative performance at the lowest levels is not needed for a stated COU. These levels are undetectable by TBS microscopy, and so it is not clear whether any amount of quantitative error would be considered clinically significant at these levels.

If the performance standard for NAAT is defined solely in comparison to the performance of TBS microscopy for detecting parasites, this would be a low bar that is likely achievable by most parasite NAATs described in the literature, regardless of their actual detection and reproducibility performance limits. If the NAATs will be used to address new COU questions, additional performance standards beyond TBS should be considered.

iv. Clinically relevant levels of parasite nucleic acids

There does not appear to be a clear consensus on what amount of parasite nucleic acids in microscopy-negative blood may be considered clinically significant, either as an indicator of peripheral blood parasite density (circulating parasites) or as an indicator of total parasite burden (circulating plus sequestered parasites).

The degree to which free RNA/DNA released from killed parasites (whether circulating or sequestered) needs to be taken into consideration will depend on factors such as the study treatment arm and the sampling time point.

These issues are applicable for NAATs that report absolute nucleic acid quantitative results traceable to an international standard, and they are similarly applicable for NAATs that report values in arbitrary units[5] to monitor relative changes over time.

4. Proposed clinically actionable biomarker threshold

   Relevant to the discussion of clinically significant nucleic acid levels, the applicant proposed a clinical biomarker threshold such that only samples with quantitative NAAT biomarker results above the threshold are considered actionable. Effectively, patients with positive biomarker NAAT results would not be given rescue treatment until the quantitative biomarker results cross the threshold level. The proposed threshold appears to have been selected partly in consideration of the analytical performance of the NAAT, and partly from the assumed significance of the predicted blood parasite density at the target threshold level (250 parasites/mL). The applicant notes that the proposed threshold will still allow subjects to be identified for rescue treatment earlier that they would be identified by TBS microscopy.

   There are a few concerns with the approach as proposed.

   a. Threshold defined as predicted parasites/mL, not copies/mL

      The applicant defined the threshold in terms of predicted parasites/mL values instead of using the NAAT biomarker copies/mL results. As noted in comment 1 above, the proposal is to qualify parasite rRNA/rDNA as a biomarker and to use quantitative NAATs calibrated to copy number, and so any discussions of clinically actionable threshold levels should be defined in terms of biomarker copy number, not predicted parasites/mL.

      If CDER/CBER entertains the use of clinical threshold biomarker levels, the predicted parasite equivalents/mL values could be used to help *justify* a clinical threshold for rRNA/rDNA biomarker levels, keeping the following points in mind:

      i. A justification that is tied to a specific predicted parasite/mL density would first require a justification for why that predicted parasite density is clinically significant.

      ii. Since the extrapolation of parasites/mL values relies on a mathematical conversion factor that will change from one NAAT instrument platform to the next, the same reported biomarker copy/mL value could have different interpretations in different studies.

[5] Absolute quantitation of nucleic acids in relation to a traceable standard may not be a strict requirement for a monitoring COU claim. CHMI involves serial testing of the same individual over time with the same test, so CDER/CBER may determine that is acceptable to monitor relative trends of biomarker concentration over time in arbitrary units, depending on the COU.

b. Significance and support for the proposed 250 parasite/mL actionable threshold

The applicant investigated various threshold values by re-analyzing a small number of serially tested subjects (n=66) from prior CHMI studies. The number of practical observations available for evaluation is smaller once the NAAT positive subjects are stratified among the different treatment arms and study protocols.

The applicant's main justifications for the selected clinically actionable threshold appear to be derived from analyzing biomarker NAAT time to positivity (TTP) results compared to either TBS microscopy TTP or symptom onset. If all test results are averaged together among all the clinical studies and treatment groups, the NAAT results do show positivity earlier than either microscopy or symptoms, regardless of whether the proposed biomarker threshold was used (Table 2).

The implication is that any threshold value that delays clinical action should be acceptable if NAAT positivity still occurs at some time point earlier than TBS microscopy positivity or symptom onset.

Table 2. Mean differences between NAAT time to positivity (TTP), TBS TTP, and symptoms[6]

	Mean difference between TBS TTP and NAAT TTP (days [95% CI])	**Mean difference between sympt. onset and NAAT TTP (days [95% CI])**
Any Positive NAAT	3.4 [3.0, 3.8]	2.2 [1.9, 2.6]
NAAT > 250 threshold	2.2 [1.9, 2.6]	1.0 [0.5, 1.4]

The applicant also proposes that 250 parasites/mL is clinically relevant because it is consistent with the number of parasites that has been predicted to be released in the first wave of blood stage parasites emerging from the liver. This argument is not very strong support on its own for why other parasite densities would not also be considered clinically relevant.

Ideally, the applicant should provide evidence for why biomarker levels (copies/mL) below a selected threshold are not indicative of clinically actionable blood stage infection.

Lastly, CDER/CBER should keep in mind that if TBS microscopy is discontinued in CHMI in favor of NAAT, the use of a biomarker NAAT threshold will affect how subjects are assessed for partial protection, and it will affect the criteria used to monitor patients after they have received rescue therapy. TBS microscopy results were previously used for these assessments.

In the clinical line data from the submission, positive TBS microscopy results became negative on average by 1.5 days (range 1 - 4 days) after rescue

[6] Total results combined among all clinical studies and treatment groups.

treatment. In contrast, positive NAAT results required an average of 6.4 days (range 1-21 days) to become negative after rescue treatment.

c. Distinction between "no biomarker detected" vs. "biomarker detected below a clinically actionable threshold"

It is important to clearly distinguish those samples with no NAAT-detectable biomarker levels from those samples with detectable biomarker levels that are below a threshold set for clinical significance. It would not be appropriate for example to refer to a sample that is positive by NAAT for parasite nucleic acids at estimated 10,000 – 100,000 copy/mL levels as "biomarker negative" simply because it is below a proposed actionable threshold level.[7] This sample would be qualitatively biomarker positive, with quantitative results reported below the proposed actionable threshold. A separate justification would be needed to support why these biomarker levels are not clinically significant or actionable.

For parasite nucleic acids detected in CHMI pre-screened negative subjects, any amount detected should be considered positive by the assay, unless there is some concern that this positive result does not reflect the true biomarker status of the sample (e.g., cross contamination, carry-over contamination, or cross reactivity with another pathogen).

5. Proposed assay cutoff

On a related issue that is separate from the discussion of a clinically actionable biomarker threshold, the applicant also proposes an assay cutoff for the biomarker NAAT. Using the proposed cutoff, NAAT results that detect the biomarker (i.e., Ct values detected < 45), but that predict parasite densities less than 10 parasites/mL are reported as NAAT negative.[8]

The justification for disregarding samples with positive Ct values is not well defined. The applicant proposes that cross reactivity and cross contamination events that typically raise concerns of false positive results for "open system" NAATs will be rare for their test. Nevertheless, there should be some rationale for questioning a low Ct value (e.g., observing an atypical amplification curve) beyond the predicted number of intact parasite equivalents. If concerns are excluded that the NAAT results are false positive and not reflective of the truth of the sample, it is not clear why positive Ct results would be disregarded, even when target levels are considered too low to be quantitated by the test.

[7] Given the proposed actionable threshold of >250 parasite equivalents/mL, and assuming the proposed conversion factor for extrapolating parasites/mL from the 3rd gen test results, a sample positive for 740,000 biomarker copies/mL would predict 100 parasite equivalents/mL (i.e., NAAT positive below the 250 parasite/mL actionable threshold).

[8] Note: Using the proposed assay cutoff, the sponsor censored the predicted parasite/mL values in the clinical line data. Parasite density values predicted below 10 parasites/mL are reported as "0 (the number < 10)", and values predicted between 10 and 19 parasites/mL are reported as "19 parasites/mL". The copies/mL values and raw Ct values appear to be reported without censoring.

For this review memo, analyses of NAAT analytical and clinical performance were conducted by counting any positive NAAT result (biomarker copies/mL and Ct) as positive, without regard to a separate assay cut-off (e.g., only positive for biomarker levels > 10 parasite equivalents/mL) or a separate clinically actionable threshold (e.g., only positive for biomarker levels > 250 parasite equivalents/mL).

Biomarker NAATs described in the submission

6. Test generations

   The submission described quantitative real-time reverse transcriptase polymerase chain reaction (RT-PCR) tests that were designed to detect parasite rRNA/rDNA in blood samples. The tests do not distinguish between parasite rRNA and rDNA in the sample. However, since rRNA copies are expected to outnumber rDNA copies in each parasite by approximately 1,000-fold (and rRNA copies per parasite are expected to increase to a certain extent as parasites mature), it is assumed that the rRNA component will represent most of the test signal.

   Over time, the applicant modified the test primers/probes, methods, protocols, and instrument platforms to produce three generations of test. While the 3rd generation test is most recent and likely to be used moving forward, both 2nd and 3rd generation tests were used in the clinical studies in the submission (Table 3). Method protocols and summary analytical data were provided for both test generations.

   Table 3. Use of 2nd and 3rd Generation RT-PCRs in Clinical Studies

RT-PCR	Clinical Study	Total # Subjects
2nd Gen	MC-003	29
3rd Gen	MC-001	6
	CVAC-PYR	21
	MC-004	10
Grand Total		**66**

7. Design differences between the 2nd and 3rd generation tests

   Some of the key design differences between the 2nd and 3rd generation NAATs are listed in Table 4. One important difference is that while the 2nd generation test reports biomarker results in a single channel for *P. falciparum* only, the 3rd generation test reports results in two channels.

   One 3rd generation test channel is used for a *P. falciparum*-specific primer/probe set, and the other channel is used for a pan-*Plasmodium* primer/probe set that is intended to detect all *Plasmodium* species without distinguishing among them. This test change avoids a potential issue with the 2nd generation test in which targets from other *Plasmodium* parasites could competitively inhibit the *P. falciparum* test signal. In the

3rd generation test, the pan-*Plasmodium* reagents are used for all biomarker detection and quantitation performance assessments.[9]

The 2nd and 3rd generation tests used different standard curve equations for calculating biomarker copies/mL from Ct values, and the applicant derived different conversion factors to estimate parasite equivalents/mL (see comments 17 - 19).

Table 4. Design Differences between 2nd and 3rd Generation NAATs

Characteristic	**2nd Generation NAAT**	**3rd Generation NAAT**
Target	*P.f.*-specific rRNA/rDNA only	Separate channels for *P.f.*-specific and pan-*Plasmodium* rRNA/rDNA
Primers/probes	*P.f.*-specific probe combined with pan-*Plasmodium* primers	New primers for both *P.f.*-specific and pan-*Plasmodium* channels. New pan-*Plasmodium* probe. Apparently uses the same *P.f.*-specific probe that was used in 2nd generation test.
Internal control	Synthetic new-fly control RNA spiked into each sample	Endogenous human gene detected as internal control
RT-PCR kit	AgPath-ID One-step RT-PCR kit	Bioline SensiFAST RT-PCR kit
Thermocycling program	~ 3 hour run time	~ 1.5 hour run time
Test-specific conversion factor to extrapolate parasites/mL	3,500	7,400

8. Analytical comparison of 2nd generation and 3rd generation performance

Head-to-head comparisons of the 2nd and 3rd generation tests were shown with a limited number of samples in the LoD (submission pg. 43) and characterized External Quality Assurance (EQA) (submission pg. 50) studies. These studies ran both tests on 10 – 20 replicates each of whole parasite samples at different parasite densities.

Qualitatively, both tests were positive 100% of the time at parasite densities of 100 parasites/mL and above. Below these densities, the 3rd generation test had a hit rate that was similar to the 2nd generation test, with slightly higher point estimates (Table 5).

Table 5. Hit Rates of 2nd Gen and 3rd Gen Tests on the Same Samples

	LoD Study 50 paras/mL	**EQA Study 60 paras/mL**
2nd Gen	18/20 (90%)	11/15 (73%)
3rd Gen	19/20 (95%)	17/21 (81%)

[9] Note: the distinction between pan-*Plasmodium* and *P. falciparum* detection is typically less of a concern for CHMI studies in which the subjects are pre-screened before enrollment and subsequently infected with a well-characterized laboratory strain of *P. falciparum*. In this regard, a single channel NAAT that detects either pan-*Plasmodium* or *P. falciparum* alone could also be acceptable for CHMI.

The applicant notes that if the estimated parasites/mL are calculated for each test using the test-specific conversion factors, the agreement for EQA samples with both test results appears to be good (submission pg. 51, Figure 17). However, the actual quantitative agreement of biomarker copy/mL result is not as high between the two tests, and so these experiments would more appropriately be described as comparisons of the different conversion factors.

CDER/CBER should keep in mind that it appears that the same EQA samples used to establish the 3rd generation test conversion factor were used to evaluate the conversion factors here. Ideally the assessment sample set should be independent of the set used to establish the conversion factor.

As noted for the other analytical studies, the predicted parasite/mL values can be used by the applicant (or by CDER/CBER) as part of the justification for why performance differences between the two test generations are not considered clinically significant.

Analytical study comments

9. Limit of detection (analytical sensitivity)

    a. Throughout the submission, the applicant proposes that the ability to detect 20 parasites/mL is clinically relevant because it would allow the detection of a single parasite in the 50 µL blood sample volume used by the NAAT. As noted, it would not be appropriate to assume that only intact parasites are detected by a biomarker NAAT in a stochastic manner. Free rRNA/rDNA could be released from parasites that lyse either within the patient or after sample collection during blood storage [10]. The LoD should be defined in terms of the qualitative (positive/negative) detection of the biomarker by counting any positive Ct values as positive test results.

    b. The pan-*Plasmodium* channel detected 95% (19/20) of replicates of parasites at nominal concentrations of 50 and 20 parasites/mL. Probit analysis indicated that the LoD was 27 parasites/mL of blood [95% CI 19, 56].

        Although the LoD study showed a 95% detection rate at 50 and 20 parasite/mL concentrations, in the separate EQA study, the pan-*Plasmodium* channel only detected 81% (17/21) of samples at 60 parasites/mL. This casts doubt on the LoD study claim that the LoD with whole parasites is 20 parasites/mL.

        Whole parasite samples at 100 parasites/mL were detected 100% (10/10) in the LoD study, and it appears that lysed equivalents of $2x10^2$ parasites/mL were detected 100% of the time in other analytical studies. The LoD should be determined by the results that can be observed experimentally. Based on the experimental data that was presented in the submission, and given the lack of samples tested between 60 and 100 parasites/mL, the whole parasite LoD could

[10] The 2-3 mL EDTA blood samples may be stored at room temperature or 4 °C, and may be held up to 48 hours prior to lysis buffer treatment.

more conservatively be set at 100 parasites/mL.

However, as noted previously, biological and experimental variability make it extremely difficult to determine the true parasite number at levels so far below the LoD of TBS microscopy. Moreover, it is not clear if the difference between LoDs of 50 or 100 parasites/mL would be considered clinically significant for the COU as currently described. Without a pre-stated clinical relevance, these difficult-to-prove LoD claims may be considered academic, and so questions about the claimed LoD value in this range should not hinder qualification. If the sensitivity of TBS microscopy is used as the performance metric for comparison, it is likely that far less sensitive NAATs could also be considered acceptable for a COU described simply to replace TBS microscopy in CHMI.

c. The applicant also estimated LoD by testing Armored RNA calibrators. The applicant tested the number of rRNA copies that would be predicted for 20 and 7 parasite equivalents/mL of blood using the assay conversion factor. For both sample concentrations, 100% of the replicates were detected by the test (21/21 and 20/20, respectively). In contrast, in the separate EQA study that tested whole parasites, samples at 6 parasites/mL were detected only 14% of the time (3/21).

   In general, tested materials should replicate clinical test conditions as closely as possible. RNA calibrators are typically not considered ideal materials for establishing test performance, as there is a concern that calibrator testing will not reflect performance with clinical samples (commutability), but may instead overestimate test performance. The wide performance discrepancy between calibrator testing and whole parasite EQA testing at similarly targeted levels (predicted 7 parasite equivalents/mL, and nominal 6 parasites/mL) highlights these concerns about the commutability of calibrator testing results.

   If CDER/CBER determine that it is useful to also calculate LoD using Armored RNA calibrators, the Armored RNA LoD results should be reported separately from whole parasite LoD results.

10. Limit of quantitation

The LoQ analysis was conducted using the two Armored RNA calibrator samples that were tested in the LoD study.

a. Armored RNA calibrators are not ideal materials for establishing test performance (9.c) due to questions of commutability. Since there is no international consensus standard for *Plasmodium* rRNA/rDNA that is based on a material more closely replicating clinical samples, RNA calibrators were used to assess NAAT biomarker LoQ in samples with defined biomarker copy values.

b. The LoQ of the biomarker NAAT should be defined in terms of biomarker copies/mL as discussed (19.c), not estimated parasites/mL as reported by the applicant. Parasites are not the test output, nor were actual parasites tested in this study. Log_{10} transformations of copy/mL values are acceptable.

Also, performance should be evaluated in terms of copies/mL of the blood sample, not in terms of copies/mL of the intermediary lysate step. The biomarker status of the patient whole blood sample, not the lysate, should be the chief concern for performance. Values were re-calculated by the reviewer for copies/mL of blood below (Table 6).[11]

*Table 6. LoQ of Armored RNA at 148,400 (5.17 log10) copies/mL **Blood** (est. 20 par/mL blood)*

Mean	**5.13**
Std. deviation	**0.20**
Bias = mean - nominal	**0.04**
TE = Bias + 2SD	0.45
[12]TE = SQRT (SD^2 + Bias^2)	0.21

The Armored RNA sample predicted to be equivalent to 20 parasites/mL met the predefined acceptance criteria after recalculation to biomarker copies/mL of blood. As a reminder, these are statistically defined acceptance criteria. CDER/CBER can decide what level of uncertainty is considered clinically acceptable for a COU, even if they exceed these statistical bounds.

If CDER/CBER entertain the use of RNA calibrators to estimate LoQ, it would still be necessary to first determine what amount of error is acceptable for biomarker quantitation in the COU.

Under the current COU, the clinical need for quantifying biomarker concentrations with very high accuracy at these low levels (or the clinical relevance of distinguishing between 20 vs. 100 parasite equivalents/mL) is not clear. Tests with higher reported LoQs (e.g., consistent with predicted parasite equivalents/mL of 100 or higher) would likely also be acceptable for the COU as currently defined.

11. Precision

The applicant conducted precision studies for the biomarker NAAT using both lysed whole parasite and Armored RNA calibrator preparations. For the whole parasite testing, %CV results were reported for estimated parasites/mL of blood values, and the nominal biomarker copy/mL values were estimated. In contrast, for the Armored RNA testing, %CV results were reported for copies/mL of lysate values, and the nominal parasite/mL values were estimated.

Neither approach is correct regarding the units. The appropriate units for describing biomarker NAAT quantitative performance are log10 biomarker copies/mL of blood, regardless of the tested material.

[11] Re-calculations should be confirmed independently.
[12] Note: TE equation #2 as presented in the current submission was corrected per CLSI EP17-A2.

The precision study results were recalculated to log10 biomarker copies/mL of blood using the line data and Excel spreadsheets provided. The results as reported by the applicant are shown in Table 7 and Table 9 for whole parasite testing and Armored RNA testing, respectively. The corresponding results recalculated to copies/mL of blood are shown in Table 8 and Table 10, respectively.

The %CVs remained as low or lower after recalculation to copies/mL of blood for both whole parasite and Armored RNA testing, with increased variability at the lowest concentrations in both scenarios. There are several potential sources for variability with these low level Armored RNA samples, such as pipetting and sampling errors that are less apparent at higher concentrations, variations in the Armored RNA material itself, or assay-specific challenges at these low levels.

The applicant indicated that the NAAT internal control reagents may compete with the *Plasmodium* reagents when parasite rRNA is less abundant than the internal control target in the sample, and this may contribute to assay variability at low target levels. Also, as discussed in the OCP review, these biomarker concentrations are far below the lowest point tested during periodic standard curve recalibrations.

As discussed for LoQ, the clinical relevance of this variability at the lowest range of the assay is not clear, and so this observation would not preclude biomarker qualification under the current COU.

Table 7. Precision - Whole Parasite – Reported Log10 Est. Parasites/mL Blood

	Samples per run	Total runs	Nominal parasites / mL blood	Nominal Predicted log10 Biomarker copies / mL blood	Pan-*Plasmodium* channel Est. Parasites/mL blood	
					Intra-assay %CV within run (95% CI)	Inter-assay %CV within lab (95% CI)
High	2	20	$4x10^5$	9.47	1.34% (1.03-1.94%)	2.74% (2.24-3.51%)
Mid	2	6	$2x10^3$	7.17	2.62% (1.69-5.78%)	5.52% (4.33-7.61%)
Low	2	20	$1x10^2$	5.87	6.63% (5.08-9.58%)	10.17% (8.33-13.06%)

*Table 8. Precision - Whole Parasite – Reported Log10 Biomarker copies/mL Blood - **Recalculated***

	Samples per run	Total runs	Nominal parasites/ mL blood	Nominal Predicted log10 Biomarker copies / mL blood	Pan-*Plasmodium* channel Biomarker copies/ mL blood	
					Intra-assay %CV within run (95% CI)	Inter-assay %CV within lab (95% CI)
High	2	20	$4x10^5$	9.47	0.79% (0.60%-1.15%)	1.60% (1.31%-2.06%)
Mid	2	6	$2x10^3$	7.17	1.19% (0.77%-2.63%)	2.51% (1.97%-3.47%)
Low	2	20	$1x10^2$	5.87	2.19% (1.67%-3.16%)	3.45% (2.82%-4.43%)

Table 9. Precision - Armored RNA – Reported Log10 Biomarker copies/mL Lysate

Samples per run	Total runs	Nominal Biomarker copies/ mL lysate	Nominal log10 Biomarker copies/ mL blood	Nominal Predicted parasites / mL blood	Pan-*Plasmodium* channel Biomarker copies/ mL blood	
					Intra-assay %CV within run (95% CI)	Inter-assay %CV within lab (95% CI)
7	3	3710	5.17	20	1.89% (1.07%-7.03%)	7.93 % (4.96-19.46 %)
6-7	3	1325	4.72	7	8.48 % (4.80-31.59%)	19.46 % (12.16-47.73%)

*Table 10. Precision - Armored RNA – Reported Log10 Biomarker copies/mL Blood - **Recalculated***

Samples per run	Total runs	Nominal Biomarker copies/ mL lysate	Nominal log10 Biomarker copies/ mL blood	Nominal Predicted parasites / mL blood	Pan-*Plasmodium* channel Biomarker copies/ mL blood	
					Intra-assay %CV within run (95% CI)	Inter-assay %CV within lab (95% CI)
7	3	3710	5.17	20	1.29% (0.73%-4.81%)	5.43% (3.39%-13.31%)
6-7	3	1325	4.72	7	5.37% (3.04-19.99%)	12.32% (7.70-30.21%)

12. Reportable range

The applicant tested serially diluted parasites over a wide range ($1x10^2$ – $1x10^7$ parasites/mL) to establish the linearity of the biomarker NAAT. The applicant reported NAAT results in log10 estimated parasites/mL.

As for other analytical studies, the reviewer recalculated NAAT results to log10 biomarker copies/mL of blood for analysis. The results were plotted against the nominal biomarker copies/mL of blood as predicted from the applicant's conversion factor and the nominal parasite density. The recalculated results were consistent with applicant's assessments of linearity over the tested range (Figure 1). The greatest variability was observed at the lowest concentration tested (estimated 5.87 log10 copies/mL, $1x10^2$ parasites/mL) in which one of the six replicates was reported with a more than 0.55 log10 (~3.5-fold) difference from nominal (Figure 2).

The trend of increasing variability observed down to $1x10^2$ parasites/mL levels is consistent with the observations of the precision study, and it is consistent with the general questions about the quality of biomarker quantitation below these levels. However, CDER/CBER may not consider biomarker quantitative differences at the lower concentrations to be clinically relevant.

Figure 1. Reportable Range Linearity

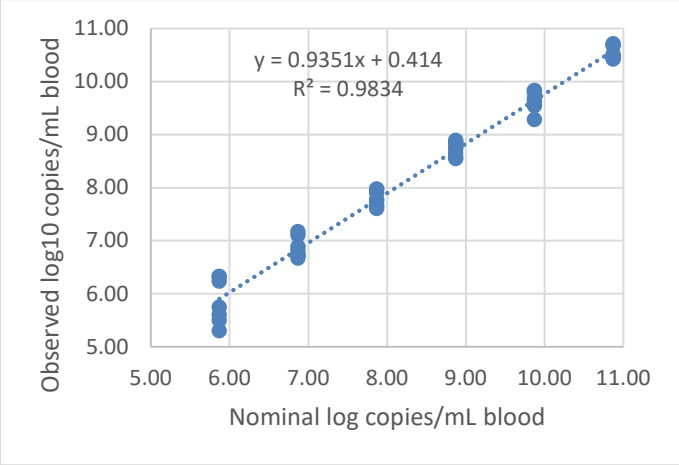
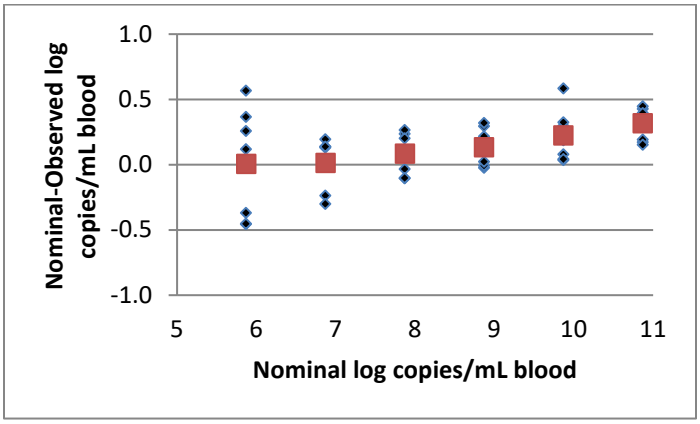


Figure 2. Reportable Range Nominal - Observed



13. Biomarker stability

The stability studies were supportive of the SOP instructions for storing unprocessed whole blood samples at room temperature for up to 12 hours and at 4 °C for up to 48 hours. Although the lysed sample stability studies tested high copy samples that did not challenge the lower range of the assay (~9.40 log10 copies/mL blood), the results were supportive of lysed sample storage up to 96 hours.

The freeze-thaw study tested single replicates of nine samples at the different freeze-thaw conditions. One sample had a three-fold decrease in copy/mL signal (within the applicant's general acceptance criteria), and one sample (~370 parasite equivalents/mL) showed more than a five-fold drop in signal. Samples frozen once are allowed for testing after consultation with the lab director and adhering to optimal collection.

14. Inclusivity

Pre-extracted DNA was tested from *P. falciparum, P. vivax, P. malariae, P. knowlesi,* and *P. ovale* (8.33E-4 ng/µL) spiked into extracted malaria-negative eluates. Clinical specimens with intact malaria RNA were tested for *P. falciparum* (n=12) and *P. vivax* (n=2). Additionally, rodent infecting strains *P. yoelii* 17XNL and *P. berghei* ANKA were tested.

This study demonstrated that the pan-*Plasmodium* reagents can detect *Plasmodium* species other than *P. falciparum,* and that they can detect parasite DNA in addition to RNA. The *P. falciparum* reagents detected *P. falciparum* strains from different sources as expected.

Inclusivity is less of a concern for CHMI studies in which subjects are pre-screened and then experimentally infected with a known *P. falciparum* strain. It appears that line data was not submitted for the inclusivity testing. For completeness of the record, line data can be requested, including Ct and copy/mL results, and estimates of the quantities that were tested in genomic equivalents/mL if that information is available.

15. Carryover

The submission described carryover testing in which a total of 207 samples (n=105 known positive and 102 known negatives) were tested interspersed within the same runs with no evidence of carryover contamination. Additionally, positive and negative controls are included with every assay run, with apparently only one possible contamination event observed in the clinical studies[13]. This is supportive data.

The line data provided for the analytical studies only flagged a total of 70 samples as carryover testing (n=35 positives and 35 negatives) instead of the 207 samples described in text. For record completeness, the applicant should update line data to flag the remaining samples and runs that were considered for carryover/contamination

[13] Possible contaminant in study MC-003, sample MC-003-025. A single positive NAAT result (34.3 Ct, 4.37 log10 copies/mL) was observed on day 12, with negative results days 6-11, and 13-19.

testing.

Other NAAT descriptive details

16. NAAT operational steps

    The sample handling and operational steps are largely the same between the 2nd generation and 3rd generation tests.
    a. A 50 µL sample of EDTA whole blood is added to two mL of lysis buffer[14].
    b. Nucleic acids are extracted from a one mL aliquot of the blood/lysis buffer mixture (~25 µL of the original blood sample) on the Abbott m2000sp instrument. The extracted nucleic acids are eluted in a final volume of 53 µL.
    c. Of the 53 µL eluate, 15 µL is run in a RT-PCR reaction on the Abbot m2000rt instrument
    d. A negative control, a low positive control, and a high positive control (contrived from lysed parasite/blood matrix) are included in each run. A maximum of 96 reactions can be performed per run.

17. NAAT result interpretation

    The RT-PCR tests generate cycle threshold (Ct) values as raw data. Log-linear standard curve values are stored from previous test runs of Armored RNA calibrators. The Armored RNA standard curves are run every six months or when new lots of reagents are introduced. The stored standard curve values are used to calculate rRNA/rDNA biomarker copies per PCR reaction from Ct values, and biomarker copies per mL of blood sample are calculated by accounting for the sample test fractions and dilutions described in the operation steps in comment 16.a (see also comment 18.a).

    For the low positive and high positive controls, the Ct values must fall within predefined ranges to be acceptable. For the negative control, no amplicon should be detected.

    As an additional calculation, test-specific conversion factors are used to extrapolate parasite equivalents/mL of blood from biomarker copy values (see comment 19).

18. Armored RNA standard curve

    The Armored RNA standard curve consists of the full-length *P. falciparum* 18S rRNA target diluted in whole blood at copy numbers defined by the manufacturer.

    The initial standard curve validation included five dilutions representing 10-fold steps from 5.3×10^7 to 5.3×10^3 and a final 2-fold dilution to 2.65×10^3 biomarker copies/mL of lysate, or 2.12×10^9 to 1.06×10^5 rRNA copies/mL of blood.

[14] For the 2nd generation test, the lysis buffer is spiked with the synthetic internal control.

a. Stored standard curve

From the SOPs provided, the stored 3rd generation standard curve equation for calculating copies per reaction appears to be

Equation 1

$$10^{[(Ct-Yint)/slope]} = x\ copies/reaction$$

with slope = -3.38 and Y-int = 41.85.

The calculations to copies/mL of lysate and copies/mL of blood are as follows:

Equation 2

$$(copies/reaction)x\ (53\ \mu L\ eluate/\ 15\ \mu L\ PCR\ template) = copies/mL\ lysate$$

Equation 3

$$(copies/mL\ lysate)x\ (1000\ \mu L\ lysate/25 \mu L\ blood) = copies/mL\ blood$$

b. Periodic curve testing

A new Armored RNA standard curve, generally consisting of three dilutions at $5.3x10^7$, $5.3x10^6$, and $5.3x10^5$ biomarker copies/mL of lysate (i.e., $2.12x10^9$ to $2.12x10^7$ copies/mL blood), is run every six months or when reagent lots are changed. From the descriptions provided in SOPs, if the new curve meets set acceptance criteria (Table 11), then the most recent unexpired stored curve values can be used to calculate biomarker copies from the Ct values.

Table 11. Curve Acceptance Criteria

Slope	-3.38 cycles/$\log_{10}$ (+/- 2SD = -3.1 to -3.6)
Y-int	41.85 cycles (+/- 2 SD = 39.44 to 44.26)
r^2	0.998 (acceptable >0.975)

As discussed in the OCP review memo, the three-point curve covers a narrow three-log copy number range and leaves an approximately two-log gap between the lowest evaluated standard point and copy numbers near the targeted LoD.

The extrapolation of copy number values at the lower end of the curve, and the re-use of previous standard curve values within set margins raise concerns about assay drift and imprecision. In the response to Information Request (8/8/17), the applicant provided information to address some of these concerns, including data from standard curves with four and five-log ranges, and plots of positive control results over time and through curve recalibrations. (See OCP review memo for analysis).

Using rough calculations of different hypothetical curves that might meet the curve acceptance criteria, it seems that a three-fold shift in calculated biomarker copy values could potentially occur between calibrations.

However, the clinical relevance of this level of possible bias is not clear when there are no clearly defined clinically relevant decision points for quantitation, and particularly when considering biomarker copy numbers that predict parasite densities far below what is detectable by microscopy. The level of acceptable quantitative bias is an issue for consideration by CDER/CBER review divisions.

19. Conversion factors used to extrapolate parasites/mL from biomarker copies/mL values

The applicant proposes test-specific conversion factors to extrapolate parasite equivalents/mL from detected biomarker copies/mL. Different conversion factors are used to account for differences in nucleic acid recovery and master mix performance between tests.

a. Potential for confusion between biomarker copies/mL of lysate and biomarker copies/mL of blood when using conversion factors

In the most straightforward description of the process, once biomarker copies/mL of blood are calculated from the standard curve, these values are divided by the conversion factor to estimate parasites/mL of blood:

Equation 4

$$\frac{Biomarker\,\frac{copies}{mL\ of\ blood}}{Conversion\ factor\,\frac{copies}{parasite}} = Estimated\,\frac{parasites}{mL\ of\ blood}$$

Note: a point of potential confusion in the submission is that the applicant sometimes describes biomarker copies/mL of lysate, not copies/mL of blood.

Biomarker copies/mL lysate values must be multiplied by 40 before dividing by the conversion factor (See 18.a), while values that have already been converted to copies/mL of blood (e.g., clinical study line data) and can be divided by the conversion factor directly.

b. Deriving the conversion factor for the 3rd generation test

For the 2nd generation and 3rd generation tests, the proposed conversion factors are 3,500 and 7,400, respectively. To derive the conversion factor for the 3rd generation test, the applicant tested 11 replicates each of characterized External Quality Assurance (EQA) panel samples containing nominal parasite concentrations of $3x10^5$ and $6x10^3$ parasites/mL, respectively. The applicant recorded the biomarker copy results for each replicate, and then calculated the conversion factor needed to achieve the nominal parasite/mL value in both the pan-*Plasmodium* and *P. falciparum*-specific channels.

i. Conversion factor combined from both pan-*Plasmodium* and *P. falciparum*-specific channels.

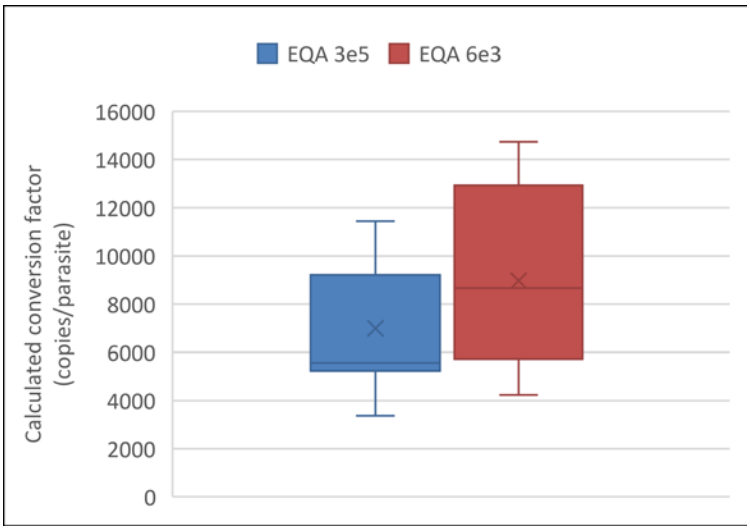
The averaged conversion factors from the pan-*Plasmodium* and *P. falciparum*-specific channels were 7,978 and 6,799, respectively. The applicant averaged these values together, and rounded up to 7,400.

Since the applicant indicates that the pan-*Plasmodium* channel results alone will be used for all detection and quantitation metrics, it may be more appropriate to use the averaged conversion factor derived from the pan-*Plasmodium* channel alone (i.e., 7,978 rounded up to 8,000). However, it is not clear if CDER/CBER would consider the resulting 7-8% difference in reported biomarker levels to be clinically relevant.

ii. Range of conversion factor values

For the pan-*Plasmodium* channel result, the averaged conversion factor value summarizes a range of values that were observed for the two EQA samples with nominal $6x10^3$ and $3x10^5$ parasite/mL densities (Figure 3). The range of values could be attributed to variability from both biological (e.g., distribution of copies per parasite) and experimental (e.g., pipetting, counting) sources.

Figure 3. Conversion Factor Values Calculated for pan-Plasmodium Channel



c. Use of estimated parasite/mL values in analytical study analyses

The applicant frequently reports analytical study results as log_{10}-transformed estimated parasites/mL. As discussed above, biomarker copies/mL values (log_{10}-transformed where appropriate) should be used when assessing the quantitative performance of the NAAT. Parasite/mL estimates may be more appropriately considered as correlations to biomarker copy numbers, not as absolute values. Predicted parasite/mL levels can be used to provide context/justifications when assessing the biomarker copy/mL results under evaluation.

Appendix I - Terminology[15,16]

Limit of Detection (LoD): the lowest concentration of analyte that can be consistently detected (typically, in ≥ 95% of samples tested under routine clinical laboratory conditions and in a defined sample type.

Limit of Quantitation (LoQ): the lowest amount of measurand that can be determined with a stated accuracy (defined as total error or as independent requirements for bias and precision) under stated experimental conditions.

Bias: the difference between the expectation of the test results and an accepted reference value. An estimate of a systematic measurement error.

Measurement precision: closeness of agreement between measured quantity values obtained by replicate measurements on the same or similar objects under specified conditions. Measurement precision is usually expressed numerically by measures of imprecision (random error) such as SD, variance, or %CV.

Total analytic error (TAE): estimate of the interval of possible analytical errors for a measurement procedure at a defined measurand content and confidence interval (typically 95% two-sided confidence interval). Includes both systematic and random errors during the examination phase of the analytical method, but does not include errors in the pre-analytic and post-examination phases.

Allowable total error (ATE): an analytical quality goal that sets a limit for both imprecision (random error) and bias (systematic error) that are tolerable in a single measurement. Reflects the maximum error in a lab result that one assumes will not lead to patient harm, alter a medical decision, or cause an adverse outcome. Ideally, the TAE is significantly less than the ATE.

The determination of ATE is ultimately a subjective decision, and depends on how the test results are used. Several sources should be used to determine ATE, including clinical outcome studies (highest standard), biological variability data, professional practice guidelines, published literature, expert opinion, statistics, and test capability.

Total error: includes combined effect of all errors throughout the total testing process, including pre-examination, examination (TAE), and post-examination phases. (No CLSI guidelines available to measure total error.)

---

[15] *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.

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

Digital Signature Concurrence Table	
Reviewer Sign-Off	Noel J. Gerald -S 2018.11.07 15:08:07 -05'00'
Branch Chief Sign-Off (Optional)	
Division Sign-Off (Optional)	Steven R. Gitterman -S 2018.11.10 20:32:32 -05'00'