September 14, 2026

Centers for Medicare & Medicaid Services
Department of Health and Human Services
Attention: CMS-3485-NC
P.O. Box 8016
Baltimore, MD 21244-8016

Re: Request for Information; Clinical Laboratory Improvement Amendments of 1988 (CLIA) Regulations; CMS-3485-NC

Dear Administrator Oz and Dr. Bhattacharya:

The Association for Diagnostics & Laboratory Medicine (ADLM) appreciates the opportunity to provide comments on the Centers for Medicare & Medicaid Services (CMS) and Centers for Disease Control and Prevention (CDC) Request for Information regarding potential updates to the Clinical Laboratory Improvement Amendments of 1988 (CLIA) regulations.

ADLM is a global scientific and medical professional organization dedicated to clinical laboratory science and its application to health care. ADLM brings together clinical laboratory professionals, physicians, research scientists, and business leaders working across clinical chemistry, molecular diagnostics, mass spectrometry, microbiology, data science, laboratory management, and other areas of laboratory medicine.

Laboratory medicine has changed substantially since the current CLIA framework was established. Advances in molecular diagnostics, next-generation sequencing (NGS), mass spectrometry, laboratory automation, digital pathology, software, and artificial intelligence (AI) have introduced new considerations for laboratory quality and oversight. ADLM supports targeted updates and guidance that reflect contemporary laboratory practice while preserving CLIA's flexible, performance-based approach.

A. Breath Testing

ADLM supports clarifying how CLIA applies to breath testing used for clinical purposes. When exhaled breath or its constituents are examined to provide information for the diagnosis, prevention, or treatment of disease or impairment, or for the assessment of health, the applicable quality requirements should reflect the nature and complexity of the testing performed. CLIA applicability should not turn solely on the physical form of the patient-derived material, the location of testing, or whether analysis occurs directly in a device or after a breath specimen is collected and transported.

CMS and CDC should also clarify when exhaled breath is considered a patient specimen or other material derived from the human body for CLIA purposes. Where clinical breath testing falls within CLIA, oversight should remain risk- and complexity-based, recognizing meaningful differences between simple, closed point-of-care systems and more complex testing that uses specialized analytical platforms or multistep interpretation.

Question 1: What breath tests are facilities performing for clinical use?

Established clinical applications include hydrogen and methane breath testing for small intestinal bacterial overgrowth and carbohydrate malabsorption, urea breath testing for Helicobacter pylori infection, and exhaled carbon monoxide testing used in smoking cessation and related clinical assessment. Breath-based testing is also an area of continued development, including volatile organic compound approaches for applications such as cancer, microbial identification, and metabolic disease.

Because the clinical maturity, intended use, and technical complexity of these tests differ substantially, CMS and CDC should avoid treating breath testing as a single uniform category. Oversight should be based on the specific test system and the risks associated with an inaccurate result.

Question 2: What methodologies and technologies do facilities use in breath testing for clinical use?

Breath testing may use gas chromatography, infrared spectroscopy, electrochemical sensors, mass spectrometry, and commercial point-of-care analyzers. Some systems measure analytes directly at the point of collection, while others require collection of a breath specimen for subsequent analysis on a separate instrument or at a different facility.

These differences have implications for validation and verification, calibration, quality control, personnel, and specimen handling. CMS and CDC should apply CLIA requirements in a manner that accounts for the methodology, intended use, degree of automation, and complexity of the examination rather than establishing a one-size-fits-all framework for all breath tests.

Question 3: What type of facilities perform breath testing for clinical use?

Clinical breath testing may be performed in hospitals, gastrointestinal clinics, and reference laboratories. In some models, collection and analysis occur at the same location; in others, a patient may provide a breath specimen at one site that is transported to another facility for analysis. CMS and CDC should clarify how CLIA responsibilities apply across these models so that responsibility for specimen collection instructions, test performance, result reporting, and quality oversight is clearly assigned.

Question 4: How do facilities collect, transport, and store clinical breath specimens? What challenges, if any, are encountered?

Breath testing can present significant preanalytic challenges. Depending on the test, results may be affected by patient preparation, timing of collection, collection technique, the collection

device, environmental contamination, specimen stability, storage conditions, and the time between collection and analysis. Tests that require administration of a substrate and serial breath collection also depend on adherence to test-specific timing and collection procedures.

These factors should be addressed through test-specific procedures that define patient preparation, specimen collection, identification, acceptability, storage, transport, stability, and

rejection criteria. The laboratory performing the examination should validate or otherwise establish appropriate requirements for the breath specimen and provide clear instructions to personnel or facilities responsible for collection.

Where a facility only collects or prepares a breath specimen and sends it to another laboratory for testing, CMS and CDC should preserve the existing distinction between specimen-collection activities and the performance of laboratory testing. A collection-only site should not automatically require a separate CLIA certificate solely because breath is the specimen type. At the same time, the CLIA-certified laboratory performing the testing should retain responsibility for establishing appropriate collection and transport requirements and for ensuring that specimens received are suitable for reliable analysis.

Overall, ADLM recommends that CMS and CDC clarify the applicability of CLIA to clinical breath testing while retaining CLIA's flexible, performance-based approach. Requirements should be appropriate to test complexity and patient risk and should support consistent quality across the preanalytic, analytic, and postanalytic phases of breath testing.

B. Laboratory Processes and Procedures

B.2. Specimen Preparation Activities and Personnel

Question 1: What activities does the laboratory consider to be part of specimen preparation?

Specimen preparation encompasses the activities necessary to ensure that an appropriate specimen is obtained, preserved, transported, processed, and prepared for testing. These activities begin before collection and continue until the specimen is ready for analysis.

Pre-collection activities may include instructions regarding fasting, medication restrictions, timing of specimen collection relative to drug administration, avoidance of known analytical interferences, and other patient preparation requirements that can affect result accuracy. Collection-related activities include selection of the appropriate collection site, needle and collection device; proper tube selection and fill volume; tourniquet application; and collection techniques appropriate to the patient population.

Following collection, specimen preparation may include transportation, accessioning, centrifugation, aliquoting, separation of serum or plasma from cellular components, tissue processing, slide preparation and staining, inoculation of culture media, DNA or RNA extraction, and other test-specific processes.

Transportation is also an important part of the preanalytic process. Specimens may move through external courier networks or internal systems such as hand delivery and pneumatic tubes. Transit time, temperature, agitation, storage conditions, and other environmental factors can affect specimen integrity and test performance.

Pediatric and neonatal testing presents additional challenges. Many laboratory automation systems are designed around adult specimen containers and volumes, while premature neonates,

newborns, infants, and small children require low-volume collection to minimize blood loss. Low-volume collection devices are not always compatible with automated laboratory systems. As a result, these specimens may require specialized centrifugation, manual aliquoting into automation-compatible tubes, relabeling, or direct loading onto analyzers. Each additional handling step increases labor requirements and creates additional opportunities for preanalytic error, specimen loss, occupational exposure, and delayed turnaround times.

CMS and CDC should recognize the full preanalytic process when considering specimen preparation requirements and allow laboratories to establish procedures appropriate to their testing systems, workflows, and patient populations.

Question 2: What is the education and experience of the personnel who perform specimen preparation activities for your laboratory?

The education and experience appropriate for personnel performing specimen preparation depend on the nature and complexity of the activities involved. Personnel may include phlebotomists, laboratory assistants, medical laboratory technicians, medical laboratory scientists, and other appropriately trained personnel.

Laboratories train personnel according to procedures that reflect their equipment, workflows, testing menu, and patient population. Training may draw on manufacturer instructions, peer-reviewed literature, professional standards, and guidance developed by organizations such as the Clinical and Laboratory Standards Institute (CLSI) and International Organization for Standardization (ISO).

CLIA should continue to allow laboratories to determine appropriate qualifications based on the complexity and risk associated with particular specimen preparation activities rather than establish a uniform personnel standard for all preanalytic functions.

Question 3: What types of training does the laboratory provide for the personnel who perform specimen preparation activities?

Training may include structured onboarding, written competency-based instruction, direct observation, and supervised hands-on practice. It should address the activities relevant to an individual's responsibilities, including specimen collection requirements, transportation procedures, specimen processing, centrifugation, aliquoting, identification of unacceptable specimens, laboratory information systems, instrumentation, safety procedures, and quality requirements.

Training should be documented and tailored to the specific duties performed.

Question 4: How does the laboratory ensure that personnel who perform specimen preparation activities remain competent?

Laboratories maintain competency through ongoing assessment appropriate to an individual's responsibilities. Assessments may include direct observation, review of work products, proficiency exercises, procedure review, corrective action monitoring, and retraining when procedures or equipment change or when performance concerns are identified.

CLIA should continue to provide laboratories flexibility to establish competency assessment processes appropriate to the activities performed and the associated risk.

B.3. Suboptimal Specimens

Question 1: What are the circumstances under which the laboratory tests suboptimal specimens?

The circumstances under which laboratories test suboptimal specimens vary by laboratory discipline, analyte, specimen type, and clinical situation. Common specimen quality concerns include hemolysis, icterus, lipemia, inadequate fill volume, clotting, improper handling, and other conditions that may affect analytical performance.

Many specimen quality issues are identified only after collection, transportation, accessioning, or analytical testing has begun. Automated interference indices and middleware systems, for example, may identify hemolysis or other interferences during or following analysis.

When a specimen is identified as suboptimal, the laboratory evaluates whether the degree of interference exceeds established criteria for the affected analyte. If the interference remains within acceptable limits, the result may be reported with an explanatory comment. If the interference could significantly affect the reliability or clinical interpretation of the result, the test may be canceled and recollection requested.

The decision to test or report a suboptimal specimen should account for the analyte, degree of interference, clinical circumstances, availability of an alternative specimen, and potential impact on patient care.

Question 2: How often does your laboratory test suboptimal specimens annually?

There is no single annual frequency that is representative of clinical laboratory practice. The prevalence of suboptimal specimens varies considerably by laboratory, patient population, clinical department, specimen type, and test.

The frequency with which results from suboptimal specimens are ultimately released also depends on analyte-specific interference thresholds established through manufacturer information, laboratory verification studies, and clinical risk assessments.

Certain clinical environments may experience higher rates of specimen quality issues because of the difficulty of specimen collection. Hemolysis, for example, may be more common in emergency departments, neonatal and pediatric intensive care units, psychiatric units, and other settings involving challenging collections.

A single national benchmark for the frequency of suboptimal specimens would therefore have limited value unless it accounts for differences in patient populations, clinical settings, specimen types, and testing platforms.

Question 3: How does the laboratory document and report results from suboptimal specimens?

Laboratories use laboratory information systems, middleware platforms, analyzer-generated flags, result comments, quality management systems, and other electronic tools to document specimen quality issues.

When testing can be completed and the result remains suitable for clinical use, the laboratory may release the result with an explanatory comment identifying the interference and its potential effect. When a specimen is unsuitable for testing or the result cannot be reliably interpreted, the laboratory documents the reason for cancellation and follows established procedures for communicating the need for recollection.

There is currently no universally adopted approach for tracking and reporting suboptimal specimens across all laboratory disciplines. Greater standardization of specimen quality indicators could improve quality management and facilitate meaningful comparison while still accounting for differences among methodologies, analytical platforms, and patient populations.

Question 4: How does the laboratory communicate with ordering providers regarding suboptimal specimens?

Communication should reflect the severity of the specimen quality issue and its potential effect on patient care.

When testing can still be performed and the interference is considered minimal, a laboratory may report the result with an interpretive comment describing the potential impact. For example, a mildly hemolyzed specimen may be reported with a comment indicating that the affected analyte could be altered by hemolysis.

When interference is severe enough to compromise result accuracy or patient safety, the test may be canceled. The ordering provider may then be notified through telephone communication, electronic messaging, or another established communication pathway so that recollection can occur.

Laboratories should retain flexibility to establish communication procedures appropriate to the clinical significance and urgency of the affected test.

Question 5: What quality assurance measures does the laboratory apply to the testing of suboptimal specimens?

Quality assurance practices may include interference studies, verification of manufacturer interference claims, establishment of laboratory-specific acceptance and reporting thresholds, monitoring of specimen rejection and cancellation rates, and investigation of recurring specimen quality problems.

Laboratories may conduct studies to evaluate the effects of hemolysis, icterus, lipemia, and other interferents across clinically relevant concentrations. These studies support appropriate reporting thresholds and decisions regarding when a result can be reliably released.

Consensus guidance such as CLSI EP07 provides standardized approaches for evaluating analytical interference. CMS and CDC should continue to recognize appropriate consensus standards and established laboratory quality management practices rather than prescribe a single approach across all analytes and testing platforms.

Question 6: What challenges does the laboratory face in managing suboptimal specimens while ensuring test result quality?

One significant challenge is that specimen quality issues may not become apparent until considerable resources have already been used for collection, transportation, accessioning, processing, and analysis. When a result ultimately cannot be released, recollection increases costs, delays testing, and creates additional burdens for patients and providers.

There is also limited standardization of performance metrics for specimen quality across laboratories, accreditation organizations, and analytical platforms. Although laboratories may monitor measures such as hemolysis rates and hemolysis-related cancellations as part of quality improvement activities, differences in patient populations, clinical departments, collection practices, and analytical systems make direct comparison difficult.

Specimen transportation presents another challenge. External courier systems may expose specimens to variations in time, temperature, agitation, and storage conditions. Internal pneumatic tube systems can expose specimens to acceleration, deceleration, and agitation that may contribute to hemolysis or other specimen quality issues. CMS and CDC should provide clearer expectations for evaluating transport systems using manufacturer information, published literature, and local validation studies appropriate to the specimen type and analyte.

Differences among instrument manufacturers also make harmonization difficult. Systems may use semi-quantitative interference indices, categorical scales, or qualitative grading systems to report hemolysis, icterus, lipemia, and other interferences. Greater standardization in the measurement and reporting of these interferences would improve interpretation and comparison across analytical platforms.

Provider education remains another important component of specimen quality. Greater understanding of how collection, transportation, hemolysis, and other preanalytic factors affect laboratory results could reduce avoidable recollections and improve interpretation of laboratory data.

B.4. Establishment and Verification of Performance Specifications

CLIA's performance-specification requirements provide an important foundation for analytical validation, but laboratory testing has evolved considerably since those requirements were developed. Technologies such as NGS, high-resolution mass spectrometry, multiplex molecular testing, digital pathology, and computational algorithms present performance considerations that do not always align with traditional analytical characteristics.

ADLM supports maintaining CLIA's performance-based framework while providing additional guidance for contemporary technologies. Validation requirements should reflect the methodology, intended use, and patient risk rather than impose the same studies across fundamentally different types of testing.

Question 1: For tests that are not FDA-cleared or approved, including modifications of FDA-cleared or approved tests, what challenges does the laboratory encounter when establishing adequate performance specifications and appropriate acceptance criteria? Specify the relevant test or procedure associated with such challenges.

A significant challenge for many laboratory-developed and modified tests is the absence of universally accepted reference methods, certified reference materials, or standardized proficiency testing programs.

This is particularly relevant for NGS panels, high-resolution liquid chromatography-tandem mass spectrometry assays, multiplex molecular infectious disease assays, cell-free DNA testing, rare disease biomarkers, and proteomic and metabolomic assays.

When an established reference method or material is unavailable, laboratories may rely on combinations of well-characterized clinical specimens, orthogonal testing methods, reference laboratories, commercial reference materials, interlaboratory comparison studies, and published consensus recommendations. Acceptance criteria may therefore need to be scientifically justified based on the methodology and intended use rather than a fixed numerical threshold.

Other challenges include limited availability of positive specimens for rare diseases and pediatric or neonatal populations, rapidly evolving genomic knowledge, lack of standardized reference materials, complex bioinformatics pipelines, software updates that require reassessment, and difficulty establishing performance when appropriate analytical reference standards are unavailable.

CMS and CDC should recognize that appropriate validation strategies differ according to the technology, intended clinical use, available reference materials, and associated patient risk.

Question 2.a: What testing methods and/or specific applications of those methods have unique performance characteristics that need to be established and are not already addressed in the CLIA regulations or guidance?

Contemporary testing methodologies that may require evaluation of performance characteristics beyond those expressly addressed in current CLIA regulations or guidance include:

- NGS, including germline testing, somatic oncology testing, minimal residual disease, pharmacogenomics, infectious disease sequencing, HLA typing, and microbial surveillance;
- Mass spectrometry, including therapeutic drug monitoring, steroid hormone testing, vitamin analysis, toxicology, newborn screening, and metabolomics;
- Digital PCR;
- Multiplex molecular assays;
- Digital pathology and AI-assisted image analysis;
- Flow cytometric minimal residual disease assays;
- Proteomic and metabolomic profiling; and
- Cell-free DNA testing.

These methodologies have different analytical principles, outputs, and intended uses. CLIA should allow laboratories to establish the performance characteristics that are scientifically appropriate to the particular application.

Question 2.b: What are those performance characteristics?

For molecular and sequencing assays, relevant performance characteristics may include limit of blank, limit of detection, limit of quantitation, variant allele frequency detection limits, coverage depth, uniformity of coverage, bioinformatics pipeline performance, variant-calling reproducibility, contamination detection, index hopping, cross-sample contamination, sequence quality metrics, and reference database version control.

For mass spectrometry, relevant characteristics may include matrix effects, ion suppression or enhancement, extraction recovery, internal standard performance, carryover, isobaric interference, chromatographic resolution, instrument drift, specimen stability, calibration model performance, and lot-to-lot reagent variability.

For multiplex assays, considerations may include cross-reactivity, target competition, analytical interference among analytes, and dynamic reporting algorithms.

For computational algorithms and software, relevant characteristics may include software version control, reproducibility, clinical decision thresholds, change control, cybersecurity considerations, and ongoing performance monitoring following software updates.

Clinical validity may also be an important performance consideration for certain laboratory-developed and modified tests, particularly where a test result is used to identify, predict, or characterize a clinical condition or outcome. Clinical validity is distinct from analytical validity

and should be evaluated in the context of the test's intended use. Laboratories should be able to rely, as appropriate, on scientifically credible existing evidence, including peer-reviewed literature, professional guidelines, curated databases, consensus recommendations, and relevant external clinical studies, rather than being expected to independently reproduce an established clinical evidence base. Where the evidence is limited, evolving, or dependent on a particular population or clinical context, laboratories should appropriately assess and document those limitations when establishing the test's intended use and reporting practices.

Not every characteristic applies to every methodology. CMS and CDC should clarify that the performance characteristics identified in § 493.1253(b)(2) are not an exhaustive checklist and that laboratories may establish characteristics appropriate to the technology, intended use, and type of output being evaluated. Consensus standards can provide more detailed, technology-specific guidance without requiring prescriptive standards in regulation that may quickly become outdated.

Question 3: How does the laboratory currently design, develop, and prepare reagents for tests developed in-house?

Clinical laboratories develop and prepare in-house reagents through established quality management systems. Depending on the reagent and application, these processes may include written procedures, vendor qualification, raw material qualification, lot acceptance testing, stability studies, storage-condition monitoring, traceability documentation, quality control monitoring, corrective action procedures, and periodic review of reagent performance.

Laboratories may prepare calibrators, quality control materials, extraction reagents, internal standard mixtures, antibody panels, molecular primers and probes, sequencing libraries, buffer systems, and other materials necessary for laboratory-developed testing.

These activities are documented within the laboratory’s quality management system and should continue to be governed by scientifically appropriate quality practices that account for differences among methodologies and testing applications.

Question 4: What types of modifications does the laboratory commonly make to FDA-cleared or approved test systems?

Clinical laboratories commonly modify FDA-cleared or approved test systems to meet patient and laboratory needs. Modifications may include:

- Use of alternative specimen types;
- Pediatric specimen-volume adaptations;
- Extension of the analytical measurement range;
- Alternative collection devices;
- Instrument platform changes;
- Alternative extraction methods;
- Updated or population-specific reference intervals;

- Reflex testing algorithms;
- Modified reporting thresholds;
- Extended specimen stability or storage conditions; and
- Automation and workflow modifications.

These changes may be necessary when the FDA-cleared intended use does not encompass important patient populations or clinical applications, including pediatric, transplant, oncology, and rare disease testing.

CMS and CDC should distinguish among modifications based on their potential effect on analytical performance and patient risk. A modification with limited potential to affect performance should not automatically trigger the same validation expectations as a change that substantially alters the performance of a test system.

CLIA should continue to focus on whether laboratories appropriately establish and verify performance rather than prescribe identical validation studies for every technology. A mass spectrometry assay, NGS panel, and AI-based image analysis system cannot be evaluated using the same validation paradigm. A flexible, risk-based framework, supported by appropriate consensus standards and technology-specific guidance, is better suited to current and emerging laboratory technologies.

B.6. Postanalytic Interpretation and Use of Artificial Intelligence (AI)

Software has long supported laboratory testing through result verification, calculations, interpretation, and reporting. AI and machine learning increasingly perform or support many of these same activities but introduce different considerations for validation, monitoring, and change control.

CLIA should distinguish between conventional deterministic software and learned models. Deterministic systems operate according to predefined rules that can generally be specified and tested directly. Learned models derive decision boundaries from data, and their performance may depend on the patient population, local workflow, and data distribution in which they are used. These differences should be reflected in how laboratories establish and monitor performance.

Question 1: How does the laboratory use software algorithms or AI tools in the postanalytic process?

Clinical laboratories use postanalytic software to verify and release results, calculate derived values, identify and route critical values, compare current and prior results, apply reflex rules, select interpretive comments, assemble report content, and support quality management.

Conventional rules-based software remains widely used. Autoverification systems release results that meet predefined criteria and route exceptions for laboratory professional review. These systems may incorporate delta checks, plausibility checks, interference checks, and reflex rules.

Deterministic calculations are also used to generate derived values, risk indices, and other reportable information.

Functions that generate or determine reportable content are part of the testing process and should be addressed within the laboratory's quality system. Software used solely for quality analytics or workflow triage presents different considerations but remains part of the laboratory's broader quality management system.

Machine learning and AI increasingly support result review, longitudinal anomaly detection, classification, and patient-specific interpretation. These tools may perform functions similar to conventional software, but their decision boundaries are learned from data rather than explicitly programmed.

Generative AI presents additional considerations. Generative models may be used to draft narrative text, populate portions of reports, or summarize prior findings. These systems can generate inaccurate or unsupported information, omit clinically important facts, or change behavior following updates to the underlying model, prompt, or knowledge base. Any generated

content incorporated into a laboratory report remains subject to the laboratory's quality and reporting requirements. Validation, implementation, and ongoing monitoring should remain under appropriate laboratory oversight.

Question 2: Under what circumstances are software functions, including certain AI tools, used for the interpretation of the results of a test? For example, NGS, histocompatibility, and pharmacogenomics testing.

Software is involved in interpretation when it transforms laboratory data into patient-specific clinical meaning or materially influences the final report. This may include classification, filtering, prioritization, risk estimation, and generation of interpretive statements. Functions that merely store, transmit, or reformat a finalized result do not present the same considerations.

Conventional deterministic software is used for functions such as sequence alignment and variant calling; variant filtering and classification; genotype-to-diplotype conversion and phenotype assignment in pharmacogenomics; HLA allele assignment, calculated panel-reactive antibody, and donor-recipient compatibility assessment; organism identification and susceptibility interpretation; and interpretation of electrophoresis and other multianalyte patterns.

Learned models increasingly augment some of these functions, including tumor classification, variant filtering and prioritization, mutational signature analysis, eplet mismatch risk estimation, prediction of antimicrobial resistance from genotype, and flow cytometry gating.

These systems fail in different ways. An error in deterministic software generally affects every case that meets the same programmed conditions and can be evaluated using cases with known expected outputs. A learned model may produce probabilistic, case-specific errors, and its performance can depend on the population being tested. A retained challenge set may demonstrate that a software pipeline continues to function, but it cannot by itself establish that a learned model remains appropriate for the current patient population.

Both deterministic and learned systems require version control. Learned models also require appropriate local evaluation before deployment and ongoing monitoring for changes that may affect performance.

Human review should be based on risk rather than required categorically for every software-assisted process. The laboratory director should determine and document the appropriate level of review based on intended use, degree of autonomy, clinical risk, and demonstrated performance. A universal manual review requirement would also be inconsistent with the longstanding use of appropriately validated autoverification systems.

CLIA oversight should focus on the laboratory's selection, verification, implementation, and ongoing use of software without duplicating FDA oversight of regulated products.

Question 3: What roles do software functions, including certain AI tools, currently play in the interpretation of histopathology slides or results?

Digital pathology illustrates the importance of evaluating software within the complete testing system and clinical workflow.

Conventional digital pathology systems include scanners, image storage and compression, transmission systems, image management software, viewing software, displays, laboratory information system interfaces, tissue detection, focus and artifact assessment, slide identification, and quantitative image analysis.

AI and machine learning systems may support scan quality assessment, detection of carcinoma or metastases, cytology prescreening, assessment of tumor cellularity, quantification of biomarkers such as HER2, PD-L1, and Ki-67, grading and classification, prognostic estimation, prediction of molecular findings from morphology, and preparation of descriptive text for pathologist review.

Several characteristics of these systems warrant particular attention.

First, some AI outputs cannot be independently confirmed through microscopic review. A prediction of molecular status from an H&E image, for example, may not be directly confirmed or refuted by morphology alone. In these circumstances, the laboratory's implementation study takes on greater importance.

Second, the interaction between the system and the pathologist is part of overall performance. Standalone algorithm accuracy does not fully characterize a tool used as decision support. The pathologist should have access to the source image, relevant quality indicators and uncertainty information, and the ability to override algorithmic output.

Third, performance may be closely linked to the imaging chain. Scanner model, stain protocol, section thickness, fixation, and other factors can affect performance. When a laboratory uses a system outside the conditions for which performance was established, the laboratory should evaluate that use accordingly.

Question 4: What methods do laboratories use to verify the performance of the software functions, including image resolution accuracy and quality and AI tools as well as the performance of computers and monitors, used with a test system?

Verification should reflect the intended use of the system, whether it is FDA-cleared or approved and used within its labeling, the degree of autonomy with which it operates, and the potential clinical consequences of an incorrect output.

Before implementation, laboratories should define the intended patient population, specimen types, inputs and outputs, users, workflow, degree of human review, foreseeable failure modes, and prospective acceptance criteria. Those criteria should reflect the intended use and potential clinical consequences of an error.

Deterministic software can generally be verified against defined specifications. Laboratories can test programmed rules at relevant boundaries and exception paths and use regression testing against cases with known expected outputs to confirm that the software continues to function as intended following changes.

Digital pathology requires evaluation of the end-to-end imaging system, including tissue capture, focus, resolution, color reproduction, compression, artifacts, data integrity, and display performance under the conditions in which the system will be used. CLIA should not incorporate hardware-specific technical specifications that may quickly become outdated. Laboratories instead should establish and document specifications appropriate to the intended use.

Learned models require a different approach because their logic cannot be exhaustively enumerated. Laboratories should evaluate performance statistically using representative cases appropriate to the intended use. Evaluation may include sensitivity, specificity, predictive values, agreement measures, calibration, repeatability, and failure or abstention rates, depending on the type of output.

Local evaluation is important because performance can be influenced by preanalytic conditions, case mix, prevalence, instrumentation, and workflow. At the same time, laboratories should not be expected to reproduce a manufacturer's entire development program. Manufacturers should provide sufficient information regarding model and software versions, intended use, input specifications, validation populations, relevant performance information, known limitations, thresholds, update procedures, and other information necessary for meaningful verification.

Ongoing monitoring should also reflect the type of software involved. Regression testing against a retained case set may be sufficient for many deterministic systems. For learned models, however, regression testing alone may not identify changes in the patient population or input distribution that affect performance.

Monitoring of learned models may therefore include periodic re-challenge with retained local reference cases and review of input quality, output distributions, failure and abstention rates, overrides, discordances, amended reports, sampled expert review, and interface failures. Changes in these measures should prompt evaluation but should not automatically be interpreted as a loss of clinical performance without further investigation.

Change control is important for both deterministic and learned systems. Relevant changes may include model version, decision thresholds, prompts, knowledge bases, scanners, stains, instruments, interfaces, infrastructure, intended use, or patient population. Software used in patient testing should not change without the laboratory's knowledge. Laboratories should be able to identify the version used and maintain an appropriate process for evaluating changes and, when necessary, rolling them back.

CMS and CDC should also clarify how the performance characteristics in § 493.1253(b)(2) apply to software outputs. Requirements developed for quantitative measurement procedures do not always map appropriately to classifiers or other software outputs. Reportable range and reference intervals, for example, may not apply. Laboratories should be able to establish performance characteristics appropriate to the type of output being evaluated.

Question 5: Are there additional technology considerations for high complexity tests, including but not limited to laboratory use of automation, laboratory use of cloud analytics, and laboratory use of artificial intelligence, that CMS and CDC should consider incorporating into the CLIA regulations?

The most immediate need is greater clarity regarding how the existing CLIA framework applies to advanced software and AI rather than creation of a separate regulatory structure.

The existing framework can address many conventional postanalytic software functions and many uses of learned models. Additional clarity is needed, however, regarding validation, local verification, lifecycle monitoring, change management, and responsibility for different aspects of system performance.

CMS and CDC should distinguish among software functions based on intended use, degree of autonomy, clinical risk, and technical characteristics. Guidance should recognize the differences between deterministic software and learned models rather than apply identical requirements to both.

Laboratories also need sufficient transparency from software and AI developers to meet their CLIA responsibilities. Laboratories should be able to identify the model or software version in use, understand the intended use and relevant limitations, review appropriate performance information, and receive notice of changes that could affect laboratory performance.

CMS and CDC should also clarify how proficiency testing and alternative assessment requirements apply when software or AI is part of the patient-testing workflow. Section 493.801(b)(1) requires proficiency testing samples to be examined in the same manner as patient specimens, but applying an AI system to proficiency testing or alternative assessment materials may not always be technically possible, particularly for image-based systems. Clear guidance would reduce inconsistent practices and survey findings.

Federal oversight should also avoid unnecessary duplication. FDA may regulate particular software products or medical devices, while CLIA governs the laboratory's responsibility for accurate and reliable testing. CLIA requirements should focus on the laboratory's selection,

verification, implementation, quality management, monitoring, and continued use of these technologies.

ADLM recommends a risk-based, technology-appropriate approach that establishes clear quality expectations while preserving the ability of laboratory directors and qualified laboratory professionals to determine scientifically appropriate methods for meeting them.

B.7. Data-Only Facilities

Question 1: What activities do data-only facilities perform to generate, or help to generate, test results and interpretations?

Laboratory testing encompasses the total testing process, including preanalytic, analytic, and postanalytic activities. Quality oversight should therefore reflect the full testing process rather

than focus solely on the physical examination of a patient specimen. Data analysis, laboratory interpretation, and reporting may occur after the initial analytical measurement but can remain integral to generating the final reportable result and supporting safe clinical decision-making.

Accordingly, the regulatory status of a data-only facility should depend primarily on the activities it performs rather than whether those activities are geographically, temporally, or organizationally separated from the laboratory that generated the underlying data. When a facility independently analyzes patient-specific laboratory data or provides specialized laboratory interpretation that generates or helps generate a test result or interpretation for clinical use, those activities are part of the total testing process and warrant appropriate laboratory oversight.

Appropriate oversight is particularly important because laboratory data are not always interchangeable across testing systems. Interpretation can be influenced by analytical methodology, analytical sensitivity and specificity, calibration and recalibration, reagent and calibrator lot changes, instrumentation, reference intervals, harmonization or lack of harmonization, and other performance characteristics. These factors may also influence the outputs of risk calculations, algorithms, and AI systems, particularly when the data used in practice differ from the methods, populations, or conditions under which an interpretive system was developed or validated.

Data-only facilities performing these functions therefore need access to sufficient information about the underlying test system, including relevant methodologies, performance characteristics, limitations, and changes that could affect interpretation. Appropriate laboratory expertise is also necessary to identify potential failure modes and determine when changes in an input, method, instrument, calibration, reagent lot, reference interval, or other testing condition require reassessment of the interpretive process.

At the same time, CLIA oversight should not be triggered merely because an entity receives or handles laboratory information. Activities limited to storing, transmitting, formatting, aggregating, or displaying completed laboratory results do not themselves constitute laboratory testing. Likewise, the ordinary use of completed laboratory results by health care professionals as

one component of clinical diagnosis and management should remain distinct from the generation of a laboratory result or specialized laboratory interpretation.

CMS and CDC should therefore focus on whether a data-only facility performs substantive analytical or interpretive activities that generate or materially contribute to a patient-specific laboratory result or laboratory interpretation. ADLM does not recommend a blanket requirement that every entity handling laboratory data obtain a separate CLIA certificate. Rather, when multiple entities contribute to the testing process, the CLIA framework should ensure that responsibility for the relevant activities is clearly assigned and that appropriate quality oversight extends across the complete process without creating unnecessary duplication.

Conclusion

ADLM appreciates CMS and CDC's efforts to evaluate whether CLIA continues to reflect contemporary laboratory practice. As the agencies consider potential updates, you should

preserve CLIA's flexible, performance-based framework while providing targeted clarification where new technologies and testing models create uncertainty. Requirements should be proportionate to the complexity and risk of the activity, support accurate and reliable testing and patient safety, and avoid unnecessary duplication or prescriptive requirements that may quickly become outdated. ADLM welcomes the opportunity to provide additional information as CMS and CDC consider updates to the CLIA regulations.

Thank you for considering ADLM's comments. If you have any questions, please email Vince Stine, PhD, ADLM's Chief Policy Officer, at vstine@myadlm.org, or Evan Fortman, MPA, ADLM's Manager of Government Affairs, at efortman@myadlm.org.

Sincerely,

Stanley F. Lo, PhD, DABCC, FADLM
ADLM President