COLLEGE of AMERICAN PATHOLOGISTS

Statement to the Clinical Laboratory Improvements Advisory Committee Meeting April 10, 2024

The role of artificial intelligence (AI) and machine learning (ML) in the clinical laboratory

The College of American Pathologists (CAP) appreciates the opportunity to provide written comments to the Clinical Laboratory Improvement Advisory Committee (CLIAC). The CAP is the world's largest organization of board-certified pathologists and leading provider of laboratory accreditation and proficiency testing programs, and continually strives to improve and advocate for excellence in the practice of pathology and laboratory medicine worldwide in service to our patients and members, practicing pathology and laboratory medicine worldwide.

The CAP believes the training and use of artificial intelligence and machine learning (AI/ML) algorithms introduces a fundamentally new kind of data analysis into the healthcare workflow that requires an appropriate regulatory framework. By virtue of their influence on pathologists and other physicians in selection of diagnoses and treatments, the outputs of these algorithms can critically impact patient care. The data patterns identified by these systems are often not exact, as there is no perfect separation of classes or predictions. Thus, there are analogies with sensitivity, specificity, and predictive value of other complex tests performed by clinical laboratories. However, in machine learning the patterns in data are identified by software and often are not explicitly revealed. Biases or subtle errors may be incorporated inadvertently into machine learning systems, and these must be identified and mitigated prior to deployment. Naturally occurring variations in healthcare context such as case mix changes, updated tests or sample preparation, or new therapies, may also change the input data profile and reduce the accuracy of a previously well-functioning machine learning system.

The CAP anticipates that in the near future AI/ML-based technologies will power highly useful applications in a broad range of medical settings including some that are performance-critical, particularly those termed Machine Learning-enabled Device Software Function (ML-DSF). For success and safe operation, the performance quality of these applications must be verified after installation and monitored over time. Performance problems may occur if there are differences in the details of local data in comparison with the data used to train the software or if the characteristics of local data drift over time. Updates to software affecting the machine learning components inherently re-define the relationship between the training and local data and require a practical and appropriate re-verification of performance to ensure safe and effective operation. Hence, ML-DSF are analogous to high complexity diagnostic testing in requiring verification at installation and robust quality control/quality assurance procedures. Because of the partial analogy of these new technologies with current diagnostic testing, the expected impact of these technologies on the practice of pathology and laboratory medicine, and the need to adhere to CLIA in the laboratory setting, the CAP has a keen interest in the regulatory approach for AI/ML technologies.

College of American Pathologists
1001 G Street, NW, Suite 425W
Washington, DC 20001
202-354-7100

CAP members have extensive expertise in providing and directing laboratory services under the Clinical Laboratory Improvement Amendments (CLIA) regulations, which require compliance with requirements through a quality system approach for overall operations and administration of the clinical laboratory. This includes the verification and validation of any new or modified tests and devices. It is important to note that there are quality practices in the laboratory specified by CLIA that are separate from operational requirements defined by a manufacturer of a medical device and approved by the FDA. While CLIA regulations are not directly applicable to other medical specialties, they may inform thinking about performance quality goals in ways that strengthen current efforts to develop AI/ML regulations and improve the consistency of their application across medical specialties. As these tools support the decision-making of providers, the role of pathologists and other specialties to interpret results must be defined.

We encourage CLIAC to work with the FDA in drafting regulation to ensure harmonization and consistency across all requirements. The FDA proposed to regulate types of AI/ML-based software as a medical device (SaMD) modifications including (1) clinical and analytical performance improvements, (2) changes in data inputs and (3) intended use of the software. The details of these kinds of modifications and the requirements for local verification and re-verification are critical and need to be better specified. Furthermore, data inputs to SaMD may be subject to variation in the real world, for example, laboratory test results can vary based on testing kit or instrument platform produced by various vendors or microscope slides produced and stained by different histology laboratories and scanned with different devices.

As such, an effective and equitable regulatory framework for machine learning in healthcare will 1) define requirements based on risk and tailored to the likelihood and magnitude of possible harm from each machine learning application, 2) require best practices for system developers including bias assessment and mitigation, 3) define appropriate best practices for verification of system performance at deployment sites, such as local laboratories, 4) define best practices for monitoring the performance of these AI/ML systems over time and mitigating performance problems that may develop, and 5) clearly assign responsibility for problems if and when they occur .

Many considerations must be addressed before regulations can be drafted. It must be determined, for example, if a SaMD will require explicit validation for use with test kits or scanning devices. If a laboratory test that is used as one of several inputs for an AI/ML predictive algorithm is changed for cost reasons to a similar test from a different vendor, would that change or invalidate a SaMD or require local re-verification? If the latter, what form of re-verification would be acceptable? In a setting where multiple algorithms are deployed, to what extent do the requirements for validation of those algorithms “lock in” methodologies and workflows for the clinical data elements upon which they depend? This kind of lock-in has the potential to reduce the organizational agility that the FDA is hoping to promote with these regulatory changes. Can general purpose validation and performance monitoring practices be defined that identify and mitigate these kinds of problems? Should data input devices such as whole slide imaging systems and chemistry and hematology analyzers be held to reproducibility standards (e.g.

color reproduction, resolution, adsorption, etc.) that keep them within some performance envelope that all SaMD manufacturers can target?

Lastly, these systems must ensure excellent performance monitoring and maintenance. Given the inherent black box nature of the advanced mathematical approaches that underpin the SaMD applications in question and the potential for drift over time there must be a robust quality control, quality assurance and quality improvement processes, including strict delta checks and a high frequency of mandatory "result" review prior to verification. Furthermore, any modification of inputs and/or intended uses, including the SaMD Pre-Specifications concept, should be viewed as an entirely new product in need of FDA approval.

Once again thank you the time to discuss the CAP's concerns and recommendations and we welcome the opportunity for further dialogue. Please contact Andrew Northup at anorthu@cap.org or 202.297.3726.

Closing,

The College of American Pathologists

College of American Pathologists
1001 G Street, NW, Suite 425W
Washington, DC 20001
202-354-7100