Guidance for OSAC Subcommittees Drafting and Updating Standards on Performance Monitoring Activities

1. *Scope*

The purpose of this document is to provide guidance to OSAC subcommittees and Scientific Area Committees (SACs) when developing standards for the administration of performance monitoring activities (PMAs). This guide includes both requirements and recommendations to be included when creating standards for proficiency testing, intra or interlaboratory comparison testing, and observation-based monitoring. It also applies to revisions of existing standards on these topics. The underlying basis of this guidance includes forensic accreditation requirements including ISO/IEC 17025, ISO/IEC 17020, and accreditation providers' supplemental requirements. This document was written from the combined multi-disciplinary experience of Forensic Service Provider (FSP) members with feedback from stakeholders including individuals from the quality, legal, and human factors communities.

Considerations are provided for administering a performance monitoring program, including test selection criteria, preparation of internal comparison tests, and evaluation of results.

2. *OSAC Mandatory Requirements for Standards Development*

OSAC's Mandatory Requirements for Standards Development (MRSD) document defines the minimum topics subcommittees and SACs shall cover when developing standards. When it comes to competency and monitoring standards, required topics that shall be addressed include proficiency testing, interlaboratory comparisons, and intralaboratory comparisons. Specifically, the MRSD requires that standards address the scope, nature, and frequency of proficiency testing, interlaboratory, and intralaboratory comparisons. This document supplements the MRSD with additional requirements and recommendations OSAC subcommittees and SACs shall consider when writing or updating standards related to PMAs and their administration.

OSAC recommends, when available and/or practicable, the use of proficiency test providers accredited in the subject matter area to ISO/IEC 17043 by an accrediting body that is a signatory to the International Laboratory Accreditation Cooperation (ILAC) Mutual Recognition Arrangement in accordance with the requirements of ISO/IEC 17011.

3. *Terminology*

3.1. Assigned value - value attributed to a particular property or characteristic of a proficiency testing item (as defined by ISO/IEC 17043:2023)

Discussion: Annex B within ISO/IEC 17043 describes how assigned values are determined. A combination of procedures is typically used based on how the PMA is created and in accordance with known limitations for the determination of assigned values used in the performance monitoring of FSPs. The listing of the most common procedures includes the use of known production information, traceable reference material values, or participant consensus.

Approved by: Forensic Science Standards Board (FSSB); FINAL Version 1, June 2, 2025

Discussion: The term "Assigned Value" is preferred over other commonly used terms such as "expected result," "right answer," "ground truth" or "consensus value" as this term better represents the complex way that PMA results should be evaluated.

3.2. Blind PMA - where the PMA item is indistinguishable from normal customer items or samples received by the FSP (e.g., PMA item is packaged and/or shipped in such a way that it remains anonymous to the FSP). (adapted from ISO/IEC 17043:2023 Annex A)

3.3. Blind PMA, Double - a type of PMA in which the participant and, at a minimum, their direct supervisor are unaware of their participation in the PMA. Coordination with staff within the FSP is still required and is done at the highest level of the organization as possible, or external to the organization if required.

3.4. Blind PMA, Single - a type of PMA in which the participant is unaware of their participation in the PMA. Staff within the FSP, including those who may directly supervise the individual, are aware of and may be assisting in the PMA.

3.5. Forensic Science Service Providers (FSSP) - an organization or individual that provides forensic science services. (OSAC Preferred term)

3.6. Forensic Service Provider (FSP) - an organization or individual that provides forensic services. (OSAC preferred term)

Note: Within this document, the use of "FSP" will represent both the FSSP and FSP terms as it applies to all PMA programs in all forensic fields.

3.7. Interlaboratory comparison - organization, performance, and evaluation of measurements or tests on the same or similar items by two or more laboratories in accordance with predetermined conditions. (OSAC Preferred Term)

3.8. Intralaboratory comparison - organization, performance, and evaluation of measurements or tests on the same or similar items within the same laboratory in accordance with predetermined conditions. (OSAC Preferred Term)

3.9. Participant - any individual who is designated to take (or be observed in) a PMA.

3.10. Performance monitoring - the ongoing process of evaluating a forensic service provider's ability to perform work. Examples include quality control measures, observation, case review, retesting, blind testing, testimony monitoring, intra- or interlaboratory comparisons, and proficiency testing. (OSAC preferred term)

3.11. Proficiency test - evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons. (ISO/IEC 17043:2023, ISO/IEC 17025:2017)

4. Significance and Use

4.1. PMAs are an essential component of a robust quality management system to ensure the validity of results. PMAs give FSPs an opportunity to assess the performance of their individual personnel, their technical procedures and equipment, and their quality system.

4.1.1. A robust quality system assesses processes through many methods including validation, training, and competency testing, as well as PMAs which together establish the capabilities and limitations of a forensic system. PMAs are ongoing evaluations of the work done within those capabilities and limitations.

4.2. A PMA program determines who and what activities need to be checked, how to check them, and at what frequency. A PMA program is a vehicle for a FSP to evaluate the validity of existing policies and procedures and ensure continual improvement in training and education for personnel, technical procedures, and quality assurance practices to facilitate growth in testing capabilities and to address emerging challenges.

4.3. Subcommittees and SACs assist FSPs by providing discipline-specific standards for PMAs. These standards provide tailored guidance for considerations related to scope, frequency, administration of PMAs, and the role of human factors in performing discipline-specific tasks.

4.4. Participation in a robust PMA program can provide additional confidence and credibility to FSPs, customers, and external parties.

5. *General Types of Performance Monitoring*

In drafting or updating standards related to performance monitoring, subcommittees, and SACs shall reference the types of PMAs and how they are applied to their disciplines.

Test Type	Design	Production, Organization, Management[1]	Participation	Considerations for Evaluation
Interlaboratory Comparison	Cover materials, methods, technologies, instruments, skills, and/or abilities used by FSP. Samples may be sourced externally, internally, or from the re-analysis of casework.	Independent 3rd party; Participating FSP	Two or more FSPs operating under separate management systems.	Use of previously established assigned values, known production information, traceable reference material values, and/or participant consensus (limited due to smaller population base, if no consensus reached, additional analysis may be required).
Proficiency Test (Form of Interlaboratory Comparison)	Cover materials, methods, technologies, instruments, skills, and/or abilities used by FSP	Independent 3rd party	Two or more FSPs operating under separate management systems. Multiple FSPs	Known production information, participant consensus, and/or traceable reference material values.
Intralaboratory Comparison	Any type and focus on any material, method, technology, instrument, knowledge, skill, and/or ability based on the needs and goals established by the FSP. Samples may be sourced externally, internally, or from the re-analysis of casework.	Within the FSP	FSP(s) operating under the same management system	Use of previously established assigned values, known production information, traceable reference material values, and/or participant consensus (limited due to smaller population base, if no consensus reached, additional analysis may be required).
Observation-based Monitoring (e.g., Testimony Monitoring)	Employed when the performance monitoring of the desired knowledge, skill, and/or ability is not available or appropriate through proficiency testing, interlaboratory comparisons or intralaboratory comparisons. Monitoring may occur during normal activities or within a mock environment.	FSP; Independent 3rd party	Any of the participation models described above.	Focus on the process and procedures followed during the observation period. In mock environments: known production information, and/or participant consensus (limited due to smaller population base).

[1] Recommend using Proficiency Test Providers that are accredited to ISO/IEC 17043.

6. Extent of Performance Monitoring Activities

In drafting standards related to PMAs, subcommittees, and SACs shall address the extent of the activity. Standards shall:

6.1. Acknowledge that PMAs are only a single part of a quality system. PMAs do not replace proactive efforts that support the FSP's quality of work such as validation, quality controls, continuing education, etc. The subcommittee should state the limitations of the PMA program inherent to their discipline.

6.2. List common discipline-specific knowledge, skills and abilities, methods, equipment, technology, and item types that are expected to be monitored for performance.

6.2.1. Discuss the acceptable level of diversity of these items to achieve comprehensive performance monitoring.

6.2.1.1. Take into consideration other elements, such as (non)routine case types submitted, difficulty level, number of samples, and frequency of utilizing implemented methods and equipment.

6.2.1.2. This should be based on the analytical or subjective characteristics of each discipline or participant's knowledge base (access to knowledge at the system).

6.2.2. Require FSPs to conduct PMAs on the types of testing they perform or are accredited to perform as listed on their Scope of Accreditation (or however named).

6.2.3. Direct FSPs to conduct PMAs only on activities for which they have validated methods and procedures.

6.3. Provide discipline-specific information for the appropriate selection or creation of materials used in non-observation-based PMAs, to include:

6.3.1. A statement that selected or created materials should be verified against the preliminary assigned value and maintain homogeneity and stability throughout the PMA.

Discussion: FSPs may choose to use casework evidence samples that have already been processed as a blind reanalysis PMA. Blind reanalysis samples must be repackaged by the FSP and provided to participants. When utilizing casework evidence samples within a PMA, there is no control of the production or composition of the test sample(s).

6.3.2. Guidance on acceptable digitization or duplication methods in the production of materials for the PMA (e.g., photographic reproductions).

6.3.3. Guidance on acceptable simulated materials or simulated creation methods to safely create performance monitoring materials (e.g., alternative testing matrices, non-hazardous pseudo materials, simulated bloodstain patterns).

6.4. Provide discipline-specific guidance on selecting appropriate components to be monitored during observation-based PMAs, both staged and casework.

6.5. Describe the intended participants (e.g., supervisors, technical leaders, and active examiners), their roles, and discipline-specific expectations in the PMA.

6.6. Provide guidance that PMAs are performed following case-like or prescribed conditions relevant to the FSP level, analysis type, or individual as the focus of the PMA.

6.6.1. Activities that are to be treated like casework include the ability to make decisions on what to test, how much to test/not test, what method to use, which instruments to use, and how many methods/instruments to use, no differently than if the samples were evidence.

6.6.2. Activities that are to be completed under prescribed conditions (i.e., non-casework scenario) include information on which methods, equipment, or technology are to be used during the completion of the activity.

6.6.3. The incorporation of consultation, verification, and technical review.

6.6.3.1. Direct FSPs to clearly communicate policies regarding the use of consultation, verification, and technical review. These policies may differ depending on the type of PMA.

6.6.3.2. Participants shall not perform verification or technical review of the same batch/group of a PMA before they complete their own analysis.

6.6.3.3. If errors are detected during verification or technical review, they shall be remediated in accordance with the FSP's quality assurance program procedures.

6.6.3.3.1. If the errors cast doubt on the performance of the participant or the FSP, additional procedures for nonconformities or corrective actions shall be followed.

6.7. Provide guidance on how results, interpretations, or opinions are reported (e.g., qualitative, quantitative, binary, use of opinion/interpretation scales, or reported as casework).

6.7.1 Guidance should not discourage or place limitations on the reporting schemes employed (i.e., the entire range of an interpretation scale is applicable within PMA activities, to include midrange responses such as inconclusive, indeterminate, or equal support).

6.8. Align with reporting standards on the OSAC Registry (including ISO/IEC 17025: 7.8.1.2) relevant to the domain of the PMA.

6.8.1. In the event of conflicting guidance related to reporting within the standards, the approach chosen should be noted by the FSP involved in the PMA.

7. Frequency of Performance Monitoring Activities

In drafting standards related to PMAs, the subcommittees, and SACs shall address the expected frequency of performance. Standards shall:

7.1. Define a period of time over which the PMA would include a representative sample of routine case and item types. For example:

7.1.1. Direct FSPs to plan a regular review of the types of cases, item types, and as applicable, analytes identified that the FSP routinely tests to ensure the selected types of PMAs are of value in evaluating laboratory and personnel performance.

7.1.2. Direct FSPs to balance the selected types of PMAs over a period of time by including case and item types the FSP does not test routinely.

7.2. Define the type of activities and number of examinations to be completed by a single FSP or individual within a defined period of time for comprehensive performance monitoring.

8. Administration of Performance Monitoring Activities

In drafting standards related to PMAs, the subcommittees, and SACs shall address the administrative expectations. Standards shall:

8.1. Ensure that PMAs are conducted with strict impartiality and confidentiality measures in place.

8.1.1. Direct FSPs to establish a process to review pressures or relationships that could affect impartiality or other human factors and define measures to eliminate or mitigate identified risks.

8.1.2. Confidentiality in the context of PMAs means that discussion between participants is not permitted prior to completion of the test. Additionally, participants need to ensure samples/data are not easily recognized or readily accessible, unique/case type case numbers are used, and participants take action to avoid observation of or interaction with others' PMA samples/data.

8.1.2.1. Participants shall direct questions or concerns about a PMA in progress to the quality management team to facilitate technical discussion as appropriate.

8.1.2.2. Additional discipline-specific measures should be defined by the subcommittee and SAC when applicable.

8.1.3. The identity of participants and any information gathered or generated during PMAs shall be kept confidential to that individual per the applicable laws, contractual agreements, or regulation activities.

8.2. Require that PMAs have a defined timeline of events that is communicated to the participants which may include: distribution of information and materials, deadline to return results, and completion of performance evaluation.

8.2.1. This requirement may not be applicable to blind or observation-based PMAs.

8.3. Address any specific topics that are required or inherent limitations for the administration of a blind PMA.

8.3.1. Documentation of the intended scope and process to keep the test blind shall be required.

8.3.2. FSPs must identify who will be notified about the existence and intended scope of the blind test.

8.3.3. Documentation of whether the PMA is single or double blind.

8.4. Require that PMA programs be reviewed annually to determine effectiveness, consistency in application, validity, and the capability or resources to meet the needs of the FSP.

8.4.1. This review may serve as another opportunity to identify gaps in an individual's training or training program, gaps in FSP policies and procedures, and the value of the selected PMAs.

9. Performance Criteria and Evaluation of Results from PMAs

In drafting standards related to PMAs, the subcommittees and SACs shall establish successful performance criteria and methods for the evaluation of the results received during the PMAs. Standards shall:

9.1. Require that the criteria for successful performance be established and communicated prior to conducting the PMA. This should be established through a combination of the expected assigned value type(s) and a reasonable expectation of results to be obtained by the FSP based on their scope of testing, policies, and validated procedures.

9.1.1. The performance criteria should include how to address:

9.1.1.1. Varying results from the same sample set due to using different methods or procedures.

9.1.1.2. Outliers through statistical or other means.

9.1.1.3. The use of materials that may later be found to be unsuitable for evaluation purposes.

9.1.1.4. Administrative errors.

9.1.2. Establish an acceptable range of results, where applicable, and cite the basis for such.

9.1.2.1. Define a specific range of results around the assigned value that will be deemed acceptable for quantitative data. These values can be expressed in either absolute terms (e.g., results must be ±2 mm from the assigned value) or in relative terms (e.g., results must be within a specific ± percentage or standard deviation of the assigned value).

9.1.2.2. Provide guidance regarding the acceptable range of responses when an interpretation scale is used.

9.1.2.2.1. Additional guidance is recommended for midrange responses (e.g., inconclusive, indeterminate, equal support).

9.1.2.2.2. A midrange response may not necessarily indicate a discrepancy, but further review by the FSP is expected if the assigned value is an identification.

9.2. Require the PMAs assigned value be determined prior to the evaluation of participant performance, to the extent possible.

9.2.1. The assigned value is determined through a combination of procedures based on how the PMA is created and known limitations. The listing of the most common procedures includes, but is not limited to, the use of known production information, traceable reference material values, or participant consensus.

9.3. Define how PMA results are evaluated, to include external criteria and the FSP's internal policies and procedures.

9.3.1. When a participant result does not meet the criteria for successful performance, direct the FSP to evaluate how that result was achieved and document the evaluation. This could include but is not limited to: amount of material used, test methods used, verification process, technical review process, and how the result was reported.

9.3.1.1. This investigation is conducted through documented procedures such as an evaluation of non-conforming work, and/or through the proficiency test providers' appeal process.

9.3.2. Particularly for interlaboratory comparisons, test design may not account for all variations in participant FSP policy, procedure, and management. PMA participants may have different requirements regarding reporting, training, and competency. Direct the FSPs to consider these differences when evaluating results of interlaboratory comparisons or similar activities involving multiple organizations.

10. Human Factors Considerations for Performance Monitoring Activities

In drafting standards related to performance monitoring, the subcommittees and SACs shall address common human factors. Standards shall:

10.1. Define the discipline-specific contextual information necessary for the determination of what methods and instruments to use during a PMA.

10.2. Provide guidance on the design and use of blind PMAs.

10.2.1. Blind PMAs provide more accurate assessments by reducing changes in participants' performance when they are aware that they are being tested (Hawthorne effect).

10.2.2. Another advantage of blind PMAs is that they assess proficiency, accuracy, reliability, and variability in a way that reflects the cumulative effects of all factors that can affect results in casework.

10.2.3. A challenge of blind PMAs is that it is difficult to produce tests that can be administered without the FSP or participant detecting that there is a difference between the PMA and an actual case submission.

10.2.3.1. Procedures shall also be established to ensure that information and materials used within blind PMAs do not enter into any official criminal databases or other legal channels that could cause adverse effects for donors of the materials.

11. Documentation and Retention

In drafting standards related to performance monitoring, the subcommittees and SACs shall address requirements for the documentation and retention of PMA records. Standards shall:

11.1. Require all documentation be generated and retained following FSP policies and any applicable laws, contractual agreements, or regulation activities.

11.1.1. PMA administrative documentation and case records are considered quality assurance records.

11.1.2. Notification to participants of results will be retained for at least the length of time the case file is retained that was worked during the PMA period.

11.2. List the information that shall be documented during the planning and administration of the PMA. This will minimally include:

11.2.1. The proficiency test plan, to include how they meet the scope and frequency requirements of the available discipline-specific standard(s).

11.2.1.1. The FSP's selection of PMAs directly impacts the FSP's opportunities to identify gaps in an individual's training or FSP's policies and procedures.

11.2.1.2. Direct FSPs to select PMAs that are determined to be technically appropriate.

11.2.2. If a PMA is sourced internally or externally.

11.2.2.1. Externally sourced performance monitoring activities shall be from providers that are accredited to ISO/IEC 17043 where available and appropriate.

11.2.2.2. Internally sourced PMAs will follow available discipline-specific standards and best practices.

11.2.3. Notification to participants of results.

11.2.4. Evaluation of PMA selection and administration.

11.3. List the information that shall be documented by the participant during the PMA. This will minimally include:

11.3.1. Information regarding the method or procedures used during the PMA, data generated from those actions, as well as the results, interpretations, opinions, or other report contents.

11.3.2. Any methods or procedures required to be used during the PMA are clearly stated and the PMA results forms provide appropriate means to return the generated information.

11.3.3. Any deviations from established procedures.

11.3.3.1. FSPs need to be aware of and consider the potential impact of such deviations when evaluating the test results.