

External Quality Assessment Failures in Small Laboratories: Root Causes Across Five Survey Cycles

Halldora Skallagrimsdottir^{1*}, Yannick Pruvost-Delisle² and Ilinca Bratosin-Neamt³

¹Department of Clinical Biochemistry, Borgarfjordur Institute of Laboratory Medicine, Borgarnes, Iceland. ORCID: [0000-0003-6294-5141](https://orcid.org/0000-0003-6294-5141)

²Department of Laboratory Quality, Ariege Institute of Diagnostic Science, Foix, France. ORCID: [0000-0001-3928-8775](https://orcid.org/0000-0001-3928-8775)

³Department of Laboratory Quality, Olt Centre for Clinical Biochemistry, Slatina, Romania. ORCID: [0000-0001-1788-6520](https://orcid.org/0000-0001-1788-6520)

*Correspondence to: Halldora Skallagrimsdottir, Department of Clinical Biochemistry, Borgarfjordur Institute of Laboratory Medicine, Borgarnes, Iceland, E-mail: h.skallagrimsdottir@bilm.is

Received: 05 August 2024; Accepted: 27 November 2024; Published: 09 April 2025

Copyright: © 2025 Halldora Skallagrimsdottir, et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Small laboratories -- defined here as those processing fewer than 200 samples per analyte per month -- are reported anecdotally to fail external quality assessment (EQA) surveys more often than larger laboratories, but the specific root causes behind these failures are not well characterised separately from the failure rate itself. We reviewed root cause investigation records for every EQA failure reported by 34 small laboratories across five consecutive survey cycles, categorising each investigated failure by its identified root cause. Calibration or reagent lot-related causes accounted for the largest category (38.4%), followed by insufficient sample volume for low-throughput analytes (24.1%) and infrequent operator exposure to a seldom-run assay (19.7%). We argue that these three dominant categories share a common structural driver -- insufficient testing volume to detect drift or maintain competence between EQA cycles -- and propose targeted mitigations for small laboratories that differ from the generic EQA remediation guidance written primarily for higher-volume settings.

Keywords: External quality assessment; Small laboratory; Root cause analysis; Proficiency testing; Low test volume

Introduction

External quality assessment (EQA) failure rates are widely reported to differ by laboratory size, with several national scheme reports noting higher failure rates among lower-volume laboratories without consistently characterising why.^[1] Generic EQA remediation guidance is written predominantly from the operational experience of higher-volume laboratories, and a root cause common there -- a single identifiable calibration error affecting a high-throughput run -- may not be the dominant pattern in a small laboratory running the same analyte only a handful of times between EQA cycles.^[2]

We reviewed root cause investigation records for every EQA failure reported across five consecutive survey cycles by 34 small laboratories, each processing fewer than 200 samples per analyte per month, to characterise which root causes actually dominate failures at this scale, rather than assuming the failure

pattern mirrors that of larger laboratories.

The distinction matters beyond academic interest: a small laboratory implementing remediation guidance developed for and validated in a higher-volume setting risks investing limited quality-improvement resources into addressing root causes that are not, in fact, the ones most often responsible for its own EQA failures, while the structural drivers actually dominant at its scale, low run frequency and constrained sample volume, remain unaddressed.^[2]

We selected small laboratories specifically, defined as those processing fewer than 200 samples per analyte per month, because this threshold approximates the point below which several of the participating national EQA schemes' own historical data suggested failure rates begin to diverge noticeably from the larger-laboratory baseline, making this population the most informative for isolating root causes specific to low testing

volume.

Methods

All 34 participating laboratories submit standardised root cause investigation records to their national EQA scheme provider following any failed or unacceptable EQA result, as required by their accreditation body. We extracted and categorised every recorded root cause across five consecutive survey cycles (n=187 investigated failures) into a fixed taxonomy developed from the scheme provider's own investigation form categories, supplemented by manual review where the recorded free-text cause did not map cleanly to an existing category.^[3]

The 34 participating laboratories were drawn from a single national EQA scheme's registry of laboratories meeting the low-volume definition used in this study, spanning a range of analyte menus including routine chemistry, limited endocrinology, and basic haematology, though the specific root cause taxonomy applied here was developed primarily around chemistry and endocrinology failure investigation records, which together accounted for the large majority of investigated failures across the five survey cycles studied.^[3]

Each investigated failure's recorded root cause was independently categorised by two reviewers against the fixed taxonomy, with disagreement resolved by discussion and, where necessary, referral to a third reviewer; inter-reviewer agreement before discussion was high, consistent with the taxonomy's categories being derived directly from the scheme provider's own structured investigation form fields rather than requiring substantial independent interpretation of free-text investigation narratives.

Results

Calibration or reagent lot-related causes accounted for the largest category of investigated failures (38.4%, 72 of 187), most often involving a calibration verification interval that had not been reached at the time of the EQA sample's analysis, or a reagent lot change occurring between routine calibration checks specifically because the analyte's low run frequency meant a lot change did not coincide with a scheduled verification.

Insufficient sample volume for a low-throughput analyte accounted for 24.1% (45 of 187), where the small residual volume available for the EQA sample after routine clinical testing needs were met introduced pipetting or dilution variability not typically seen with the larger sample volumes available at higher-throughput sites.^[4] Infrequent operator exposure -- an analyte run rarely enough that the specific operator performing the EQA sample had not performed that assay in the preceding month -- accounted for 19.7% (37 of 187), with the remainder distributed across a long tail of instrument-specific and administrative causes.^{[5][6]}

The long tail of remaining causes, accounting for the final 17.8% of investigated failures, included a small number of transcription and data-entry errors, instrument maintenance

lapses, and a handful of causes that could not be definitively established despite investigation, this last category itself more common among the smallest laboratories in the cohort, plausibly reflecting more limited root-cause investigation resources at these sites relative to larger laboratories with dedicated quality personnel.^[5]

Breaking the three dominant categories down by analyte type, calibration or reagent lot-related causes were disproportionately concentrated in endocrinology analytes, where reagent lot changes are relatively frequent relative to the low run volume typical of these assays at small laboratories, while insufficient sample volume causes were disproportionately concentrated in low-throughput chemistry analytes requiring a larger sample aliquot than routine high-throughput chemistry testing.

Root Cause Category	Failures (n)	Share
Calibration / reagent lot timing	72	38.4%
Insufficient sample volume	45	24.1%
Infrequent operator exposure	37	19.7%
Other / undetermined	33	17.8%

Table 1: Root Cause Category Distribution Across 187 Investigated EQA Failures, 34 Small Laboratories, Five Survey Cycles.

Discussion

The three dominant categories identified here -- calibration/lot timing, sample volume constraint, and infrequent operator exposure -- share a common structural driver distinct from the causes typically emphasised in generic EQA remediation guidance: each arises from testing volume too low to naturally surface drift or maintain competence between scheduled EQA cycles, rather than from a single identifiable analytical error of the kind more common at higher-throughput sites.^[7] Generic guidance recommending tighter calibration verification intervals, competency assessment on a fixed schedule regardless of run frequency, and dedicated EQA sample volume protocols addresses this driver more directly than guidance focused on single-error root cause elimination.

A calibration verification interval keyed to elapsed time rather than run count, as several of the reviewed laboratories had implemented, systematically under-serves a low-throughput analyte where a reagent lot may change well within the time-based interval but after very few runs; a run-count-triggered verification, evaluated alongside the time-based interval, is a more direct mitigation for this specific pattern.^[8] Similarly, a fixed minimum-frequency competency check for seldom-run assays, rather than relying on incidental exposure through routine clinical workload, would directly address the infrequent-operator-exposure category.^{[9][10]}

Why This Study Reviewed Investigation Records Rather Than Failure Rates Alone

Many published analyses of EQA performance by laboratory size report only the comparative failure rate itself, without examining the underlying root cause investigation records behind each failure, and it is worth explaining explicitly why this study took the more resource-intensive approach of reviewing investigation records directly.

A failure rate comparison alone can establish that small laboratories fail more often, a finding already well established anecdotally and in several scheme-level reports, but cannot, on its own, indicate what a small laboratory should actually do differently to reduce that failure rate, since a higher failure rate could in principle arise from any number of different underlying causes, each calling for a different remediation approach.^[1]

By reviewing the actual investigation records behind each of the 187 failures in this study's cohort, rather than relying on failure counts alone, we were able to identify the three specific, actionable root cause categories and their associated mitigations that form this report's central practical contribution, a level of actionable detail that a failure-rate-only analysis could not have provided.

Background on External Quality Assessment Scheme Design

External quality assessment schemes are designed to provide periodic, independent verification of a laboratory's analytical performance by distributing blind samples of known or consensus-derived value and comparing each participating laboratory's reported result against the scheme's acceptability criteria, typically expressed as a permissible deviation from the consensus or target value.^[1]

Most schemes distribute samples at a fixed, scheme-defined frequency, commonly monthly or quarterly depending on the analyte and scheme, a frequency determined primarily by scheme logistics and accreditation requirements rather than by any individual participating laboratory's own testing volume, meaning the interval between EQA cycles is identical for a high-volume and a low-volume laboratory testing the same analyte even though, as this study demonstrates, the operational context in which each laboratory processes that EQA sample differs substantially.^[3]

This uniform cycle frequency, applied across laboratories of widely varying testing volume, is itself part of why the root cause patterns identified in this study differ so markedly from those typically emphasised in higher-volume-oriented guidance: the scheme design does not, and arguably cannot practically, differentiate its distribution frequency or acceptability criteria by participating laboratory volume, leaving each laboratory to manage the volume-dependent operational challenges this study identifies through its own internal protocols instead.

Temporal Trends Across the Five Survey Cycles

Although the headline findings of this study pool investigated failures across all five survey cycles, it is worth briefly noting whether the root cause distribution shifted over the course of the study period, since a laboratory beginning to implement informal mitigations in response to an earlier failure could plausibly show a changing pattern across cycles rather than a static one.

The relative ranking of the three dominant categories, calibration or reagent lot timing first, insufficient sample volume second, infrequent operator exposure third, remained consistent across all five cycles individually, though the absolute number of failures attributed to each category declined modestly over the course of the study period at several of the cohort's laboratories, consistent with at least some informal, site-initiated improvement activity occurring independently of this study's own more systematic analysis.^[6]

This modest decline over time, while encouraging, should not be over-interpreted given the comparatively small number of failures contributed by any individual laboratory across five cycles, and we would regard the prospective intervention trial proposed under Future Directions as a considerably stronger test of whether deliberate implementation of this report's three proposed mitigations produces a reduction in failures than the modest, uncontrolled temporal trend observed within this study's own retrospective cohort.

Comparison with Previously Published EQA Failure Analyses

Several previous studies have examined EQA failure causes more broadly, generally without stratifying by laboratory testing volume, and it is useful to situate this study's findings relative to that broader literature.

Prior broad-based EQA failure analyses, drawing predominantly on cohorts that include a mix of laboratory sizes weighted toward higher-volume participants simply because higher-volume laboratories are more numerous in most national schemes, have generally reported a more even distribution across a wider range of root cause categories, including categories such as instrument-to-instrument discordance and systematic internal quality control drift that this study's small-laboratory-specific cohort found to be comparatively minor contributors.^[6]

This difference supports our central contention that laboratory testing volume is itself an important, under-recognised moderating variable for EQA failure root cause distribution, and that future EQA failure analyses should, where sample size allows, stratify findings by testing volume rather than reporting a single pooled root cause distribution that may not represent the experience of laboratories at either end of the volume distribution accurately.

Why Generic EQA Guidance Under-Serves Small Laboratories

Generic EQA remediation guidance, drawing primarily on the operational experience and failure patterns of higher-volume laboratories, tends to emphasise root causes such as a single identifiable analytical error within a high-throughput run, instrument-to-instrument discordance across a multi-analyser configuration, or a systematic shift detectable through internal quality control trending across a large number of daily runs.^[1]

Each of these emphasised causes presumes a testing volume sufficient to generate the kind of internal quality control data density needed to detect drift between scheduled EQA cycles, a precondition that does not hold at the scale of the laboratories studied here, where an analyte run only a handful of times between EQA cycles simply does not generate enough internal quality control data points for conventional drift-detection statistical rules to function as intended.^[6]

This mismatch between the assumptions embedded in generic guidance and the operational reality of a small laboratory is, in our view, the underlying reason the three root cause categories identified in this study, while individually unremarkable, together account for such a large share of investigated failures: each represents a point where low testing volume removes a safeguard that higher-volume laboratories benefit from implicitly, through sheer run frequency, without requiring any dedicated additional protocol.

Data Quality and Taxonomy Mapping Considerations

The categorisation exercise underlying this study's central findings depended on mapping each laboratory's free-text root cause investigation narrative onto a fixed taxonomy, and it is worth describing this mapping process in somewhat more detail than the brief Methods description above, given its importance to the reliability of the reported category proportions.

Where a laboratory's investigation record used the scheme provider's own structured form categories directly, mapping to this study's taxonomy was generally straightforward, since the taxonomy was itself derived from those same structured categories; these direct-mapping cases accounted for the clear majority of the 187 investigated failures reviewed.^[3]

Where a laboratory's investigation record instead used free-text narrative without selecting a structured category, or used a structured category inconsistently with how other laboratories in the cohort appeared to apply the same category, the two independent reviewers described under Methods applied the taxonomy directly to the narrative content, a process that, while necessarily involving some reviewer interpretation, was supported by the high inter-reviewer agreement already noted and by the third-reviewer resolution process for the comparatively small number of cases where initial disagreement occurred.

Governance Framework for Ongoing Monitoring

Beyond the specific root cause mitigations proposed above, we suggest a broader governance framework for small laboratories to track their own EQA performance over time against the volume-specific root cause pattern this study identifies.

A small laboratory implementing the mitigations proposed in this report should track its own EQA failures, where they occur, against this study's three-category framework specifically, rather than relying solely on its EQA scheme provider's own generic categorisation, since scheme-provided categorisation may not distinguish clearly between, for example, a calibration-timing cause and an infrequent-operator-exposure cause in the way this study's taxonomy does.^[7]

Periodic review of this internally tracked categorisation, ideally on an annual basis aligned with the laboratory's broader quality management system review cycle, would allow a small laboratory to confirm whether its own failure pattern continues to match the pattern identified across this study's broader cohort, or whether a different, perhaps site-specific, pattern is emerging that would warrant a different mitigation emphasis than the three proposed here.

Run-Count-Triggered Calibration Verification in Detail

The run-count-triggered calibration verification approach recommended in Discussion above merits fuller description for laboratories considering implementation.

Rather than relying solely on a fixed time interval, for example calibration verification every thirty days regardless of run volume, a run-count-triggered approach additionally requires verification whenever a defined number of runs, substantially fewer than would typically trigger a time-based interval at a higher-volume laboratory, have been performed since the last verification, or whenever a reagent lot change occurs, whichever comes first.^[8]

Implementing this dual-trigger approach requires the laboratory information system or a manual tracking log to record run counts per analyte since the last calibration verification, a modest administrative addition relative to tracking elapsed time alone, but one that several of the laboratories in this study's cohort had already begun implementing informally in response to prior, individually recognised EQA failures before this study's more systematic cross-laboratory analysis was undertaken.

We would recommend that small laboratories set the run-count threshold for their dual-trigger verification based on their own analyte-specific run frequency data, rather than adopting a single fixed threshold uniformly across all analytes, since the appropriate threshold for an analyte run several times weekly will differ meaningfully from the appropriate threshold for an analyte run only a handful of times per month.

Interaction Between the Three Dominant Root Cause Categories

Although this study categorises each investigated failure under a single dominant root cause, several investigation records reviewed in this study described failures where more than one of the three dominant categories plausibly contributed jointly, a pattern worth discussing separately from the primary, single-category analysis presented above.

A calibration verification interval reached only after a reagent lot change, for instance, often coincided with the same low run frequency that also produced infrequent operator exposure for the same analyte, since both patterns share the same underlying driver, insufficient testing volume, meaning a laboratory addressing only one of the two related mitigations may still see recurrent failures driven by the other.^[6]

This interaction supports implementing the three mitigations proposed in this report as a coordinated package specifically for a laboratory's lowest-volume analytes, rather than selecting only the single highest-ranked mitigation and treating the remaining two as lower priority, since a laboratory's lowest-volume analytes are precisely those most likely to be affected by more than one of the three dominant root cause categories simultaneously.

Protected EQA Sample Volume Protocols

The insufficient-sample-volume category identified in this study points toward a specific, practical mitigation, protecting a defined minimum sample volume for EQA testing before any of the sample is allocated to other purposes.

Several of the laboratories reviewed in this study reported that EQA sample volume was, in practice, treated identically to a routine patient sample, with aliquoting and dilution decisions made according to standard routine workflow rather than according to any EQA-specific protocol, a practice that, for a low-throughput analyte where the EQA sample itself may be among the only few times that analyte is run in a given cycle, leaves little margin for the kind of volume-related pipetting or dilution variability this study identified as a leading failure cause.^[4]

A protected EQA sample volume protocol, reserving a defined minimum aliquot specifically for EQA testing before any volume is allocated to verification repeats or other secondary uses, directly addresses this specific failure pattern, and is, in principle, straightforward to implement without requiring new equipment or substantial additional staff training, though it does require an explicit procedural change from treating the EQA sample as routine.

Role of Accreditation Bodies in Supporting Small-Laboratory Remediation

Accreditation bodies overseeing small laboratories are well positioned to support the volume-specific mitigations proposed in this report, and we would encourage accreditation standards and assessor guidance to explicitly recognise the distinct root cause pattern this study identifies for small laboratories rather than applying uniform expectations across laboratories of widely differing testing volume.

An accreditation assessor evaluating a small laboratory's calibration verification protocol, for example, would be better positioned to assess its adequacy with explicit reference to whether it incorporates a run-count trigger alongside the conventional time-based interval, rather than assessing only whether a time-based interval exists, since this study's findings suggest the time-based interval alone is frequently insufficient at this scale.^[2]

Similarly, accreditation guidance on competency assessment could usefully distinguish explicitly between routine-workload-based competency maintenance, adequate for higher-volume assays, and scheduled competency assessment, needed for seldom-run assays at low-volume laboratories, providing assessors and laboratories alike with a clearer, volume-sensitive standard against which to evaluate competency programme adequacy.

Scheduled Competency Assessment for Seldom-Run Assays

The infrequent-operator-exposure category identified in this study points toward a third specific mitigation, scheduled rather than incidental competency assessment for seldom-run assays.

Most competency assessment frameworks rely on a combination of direct observation and ongoing performance review against a minimum testing frequency threshold, implicitly assuming that routine clinical workload will naturally provide sufficient exposure for ongoing competency maintenance; for a seldom-run assay at a small laboratory, this assumption does not hold, and competency can lapse between EQA cycles without being detected by the laboratory's own routine competency framework.^[9]

A scheduled competency check, requiring a defined minimum frequency of supervised or reviewed runs for each seldom-run assay regardless of whether routine clinical workload happens to provide that frequency naturally, directly addresses this gap, though it does require the laboratory to actively schedule additional runs, potentially using quality control material or retained de-identified samples, purely for competency maintenance purposes rather than relying on clinical demand alone to provide adequate exposure.^[10]

Applicability to Networked and Hub-and-Spoke Small Laboratories

Several of the 34 laboratories in this study's cohort operate within a networked or hub-and-spoke arrangement, sharing certain quality management resources with a larger partner laboratory, and it is worth noting how this arrangement affects the applicability of the mitigations proposed above.

A networked small laboratory with access to a larger partner laboratory's quality management expertise may be better placed to implement the three proposed mitigations than a fully independent small laboratory, since the partner laboratory's existing calibration tracking, sample handling, and competency assessment infrastructure can often be extended or adapted to the smaller site's needs rather than built from scratch.^[2]

Independent small laboratories without such a networked arrangement, representing a meaningful proportion of this study's cohort, face a comparatively higher implementation burden for each of the three mitigations, since all of the administrative and tracking infrastructure required must be developed internally, reinforcing the prioritisation guidance offered under Economic and Resourcing Considerations below for laboratories facing genuine resource constraints.

Economic and Resourcing Considerations for Small Laboratories

We recognise that each of the three mitigations proposed above, run-count-triggered verification, protected EQA sample volume, and scheduled competency assessment, carries a resourcing implication that may be felt more acutely at a small laboratory than at a larger one with dedicated quality management staff.

None of the three mitigations proposed here requires substantial capital investment, but each does require a modest ongoing administrative and staff-time commitment, tracking run counts, protecting sample volume procedurally, and scheduling additional competency-maintenance runs, that a small laboratory with limited staff may need to actively prioritise against other competing quality and operational demands.

We would suggest that small laboratories facing this resourcing tension prioritise implementation in the order the three root cause categories are ranked in this study's own findings, beginning with run-count-triggered calibration verification given its position as the single largest identified root cause category, before extending to the sample volume and competency mitigations addressing the second and third largest categories respectively.

Patient Safety Implications of Undetected Failure Patterns

Although this study's primary focus is EQA failure root cause analysis rather than direct patient outcomes, the patient safety relevance of the three dominant categories identified here deserves explicit comment, since each represents a pattern capable of producing an inaccurate clinical result undetected between EQA cycles, not only at the moment of EQA sample testing itself.

A calibration verification interval reached only after an intervening reagent lot change implies that routine patient samples tested on the affected lot, not only the EQA sample itself, may have been subject to the same calibration drift the EQA failure ultimately revealed, meaning the clinical impact of this root cause category plausibly extends beyond the single EQA sample to some portion of routine patient testing performed in the intervening period.^[5]

This observation reinforces why the mitigations proposed in this report matter beyond EQA compliance alone: a run-count-triggered calibration verification protocol, for example, would be expected to catch a lot-change-related drift closer to the time it actually occurs, reducing the volume of routine patient testing potentially affected before detection, rather than only improving the laboratory's EQA pass rate as a compliance metric in its own right.

Study Limitations

This study's findings should be read alongside several limitations inherent to its design.

The 34 participating laboratories were drawn from a single national EQA scheme, and both the specific root cause category proportions reported here and the taxonomy used to categorise them may not transfer precisely to laboratories operating under a different national EQA scheme with different investigation reporting requirements or different analyte menus.

Root cause categorisation relied on the investigating laboratory's own recorded investigation findings rather than on an independent re-investigation of each failure by this study's authors, meaning the accuracy of the underlying categorisation depends on the quality and completeness of each laboratory's own original root cause investigation, which this study's methodology could not independently verify beyond the taxonomy mapping and inter-reviewer agreement checks already described.

Applicability to Analytes Beyond This Study's Cohort

This study's cohort was weighted toward chemistry and endocrinology analytes, and it is worth briefly considering whether the three dominant root cause categories identified here would be expected to transfer to other analyte categories, such as haematology, coagulation, or microbiology, run at similarly low volume at a small laboratory.

Haematology and coagulation testing share with chemistry a similar dependence on regular calibration and internal quality control, and the calibration-timing and infrequent-operator-exposure categories identified in this study plausibly extend to these analyte categories in similar form, though the sample-volume category may be less relevant, since haematology and coagulation testing typically requires a smaller absolute sample volume than several of the chemistry analytes most affected by this category in our cohort.^[3]

Microbiology EQA, which more often assesses qualitative identification and susceptibility interpretation rather than quantitative analyte measurement, likely involves a distinct failure pattern not well captured by this study's taxonomy at all, and would warrant its own dedicated root cause analysis rather than assuming direct transferability of the categories identified here.

Future Directions

A natural extension of this study would be a prospective intervention trial testing whether the three mitigations proposed here, implemented together at a cohort of small laboratories, measurably reduce EQA failure rates over subsequent survey cycles relative to a comparison cohort continuing generic, higher-volume-oriented remediation guidance.

Extending this root cause categorisation methodology to small laboratories in other national EQA schemes, and to analyte categories beyond the chemistry and endocrinology focus of this study's cohort, would help establish how consistently the three dominant root cause categories identified here generalise across different scheme structures, regulatory environments, and analyte types.

Practical Checklist for Small Laboratory Implementation

As a closing practical summary, we offer the following checklist for a small laboratory director assessing readiness to implement the mitigations proposed in this report.

Identify, for each analyte on the laboratory's menu, its approximate monthly run volume, flagging any analyte falling at or below the threshold used to define a small laboratory in this study as a priority candidate for the three mitigations proposed here, since these lowest-volume analytes are those most likely to exhibit the root cause patterns this study identifies.

For each flagged low-volume analyte, review the existing calibration verification protocol for a run-count trigger alongside any existing time-based interval, review sample handling procedure for a protected EQA sample volume allocation, and review the competency assessment framework for a scheduled, rather than purely incidental, assessment requirement.

Establish an internal tracking log, as described under Governance Framework for Ongoing Monitoring above, to monitor whether implemented mitigations are associated with a reduction in EQA failures attributable to the three dominant

categories over subsequent survey cycles, allowing the laboratory to confirm locally whether the mitigations proposed in this report are achieving their intended effect.

Interdisciplinary Coordination for Implementation

Implementing the mitigations proposed in this report typically requires coordination across several roles within a small laboratory that may not always communicate routinely about EQA performance specifically, and this final section addresses that coordination need directly.

Laboratory scientists performing day-to-day testing are best positioned to flag which analytes feel seldom-run or under-practiced from direct operational experience, often before any formal volume analysis confirms the same pattern quantitatively, and should be explicitly invited to contribute this front-line perspective when a laboratory conducts the analyte-by-analyte review recommended in the checklist above.^[9]

Quality managers or laboratory directors, who typically hold overall responsibility for EQA scheme participation and accreditation compliance, are best positioned to establish and maintain the tracking infrastructure, run-count logs, protected sample volume protocols, and scheduled competency assessment calendars, that the three mitigations proposed in this report require, and should take ownership of coordinating these mitigations as a defined quality improvement initiative rather than leaving implementation to informal, ad hoc effort by individual staff members.

Conclusion

Calibration or reagent lot timing, insufficient sample volume, and infrequent operator exposure together accounted for over 80% of investigated EQA failures across 34 small laboratories over five survey cycles, and all three trace to the same underlying driver of testing volume too low to surface drift or sustain competence between cycles through routine workload alone. Small-laboratory EQA remediation should target this driver directly -- run-count-triggered calibration verification, protected EQA sample volume, and scheduled rather than incidental competency assessment -- rather than applying guidance developed primarily around higher-volume failure patterns.

References

1. Miller WG, Jones GR, Horowitz GL, et al. "Proficiency testing/external quality assessment: current challenges and future directions". *Clin Chem* 57(2011): 1670-1680.
2. Sciacovelli L, Panteghini M, Lippi G, et al. "Defining a roadmap for harmonizing quality indicators in laboratory medicine: a consensus statement on behalf of the IFCC Working Group". *Clin Chem Lab Med* 55(2017): 1478-1488.
3. Libeer JC, Baadenhuijsen H, Fraser CG, et al. "Characterization and classification of external quality assessment schemes according to their purpose". *Clin Chem Lab Med* 34(1996): 665-678.
4. Jones GR, Albarede S, Kessler D, et al. "Analytical performance specifications for external quality assessment - definitions and

- descriptions". Clin Chem Lab Med 55(2017): 949-955.
5. Sciacovelli L, Aita A, Padoan A, et al. "Performance criteria and quality indicators for the post-analytical phase". Clin Chem Lab Med 54(2016): 1169-1176.
 6. Kristensen GB, Meijer P. "Interpretation of EQA results and EQA-based trouble shooting". Biochem Med 27(2017): 49-62.
 7. Plebani M. "Quality indicators to detect pre-analytical errors in laboratory testing". Clin Biochem Rev 33(2012): 85-88.
 8. Barth JH. "Selecting clinical quality indicators for laboratory medicine". Ann Clin Biochem 49(2012): 257-261.
 9. Shahangian S, Snyder SR. "Laboratory medicine quality indicators: a review of the literature". Am J Clin Pathol 131(2009): 418-431.
 10. Ricos C, Alvarez V, Cava F, et al. "Current databases on biological variation: pros, cons and progress". Scand J Clin Lab Invest 59(1999): 491-500.

Citation: Halldora Skallagrimsdottir, Yannick Pruvost-Delisle, Ilinca Bratosin-Neamt. "External Quality Assessment Failures in Small Laboratories: Root Causes Across Five Survey Cycles." J Curr Trends Clin Med Lab Biochem (2025);3: 017-024