

Investigating the Performance of Unacceptable Results in Proficiency Testing Among Medical Laboratories

Abdalla Eltoun Ali^{1,2*}, Taha Hassan¹, Maram Elamim¹, Alneil M. Hamza³

¹Raidx Medical Laboratory, Yanbu St, Dhahrat Laban Dist. Riyadh Saudia Arabia.

²Alzaiem Alazhari University-faculty of medical laboratory science- Clinical Biochemistry Department, Sudan.

³Clinical Laboratories Sciences Department, College of Applied Medical Sciences, AlJouf University, Saudi Arabia.

*Corresponding author: Abdalla Eltoun Ali.

Abstract

Proficiency testing (PT) is crucial to ensuring the quality and accuracy of medical laboratory testing. This review article investigates the performance of unacceptable results in PT among medical laboratories. Unacceptable PT results indicate potential issues in laboratory performance, which can have significant implications for patient care. This review examines the factors contributing to unacceptable PT results, including pre-analytical, analytical, and post-analytical errors. It also explores the impact of unacceptable results on laboratory quality management and patient outcomes. Strategies for reducing unacceptable PT results and improving overall laboratory performance are discussed. Understanding and addressing the factors contributing to unacceptable PT results are essential for maintaining high laboratory standards and ensuring the reliability of laboratory test results.

Keywords: proficiency testing; unacceptable results; laboratory errors

Introduction

When a laboratory announces a PT scheme, they commonly feel confident and even relaxed when preparing their systems and starting testing samples. Some laboratories may struggle to meet deadlines and rush sample examinations due to current workloads and time constraints. After obtaining their PT scores, it could be quite a shock for those confident to learn that their test results were not as accurate as they thought. This can then lead to actions on how and why mistakes were made and what systems can be changed to correct these mistakes. PT is a learning experience for medical laboratory scientists; often, it is from mistakes that a person can learn the most. PT results allow a laboratory to see what mistakes were made and which analytes. Mistakes can range from simple clerical errors to problems with sample preparation and analyte testing [1-3]. PT samples provide an alternative to checking test results' accuracy, compared to patient samples, which is a far riskier and less ethical practice. PT mistakes should be investigated so changes can be made to improve and prevent similar mistakes from being made in the future. PT programs can also often identify a minority of laboratories providing unreliable test results. This could be due to financial constraints and a lack of

resources or a lack of qualified staff and training. In turn, it may be in the best interests of these laboratories to reconsider whether they should continue testing these analytes and even consider sending samples to a more reliable external source [1,3].

A proficiency testing (PT) program is a tool used to assess the quality of laboratory test results by comparing them to a "gold standard." This is done to monitor the ongoing capability of a laboratory to deliver a valid result and to detect any changes to the systems being used. In clinical chemistry, a PT program simulates patient samples, which can be either internal (samples are tested on-site) or external (samples are sent to another laboratory). Accurate and reliable reporting of patient test results is essential for modern healthcare. Any errors made during clinical testing can result in misdiagnosis and incorrect treatments and jeopardize patient safety. Therefore, medical laboratory scientists need to participate in PT programs to ensure their test results are reliable and that any mistakes can be identified and corrected [1,3,4].

Importance of Proficiency Testing

The importance of PT has become evident to government bodies, which now use PT participation to determine eligibility for participation in health care schemes. PT has also been endorsed by the European Union's Directive on using standards in the pursuit of standardization to assist laboratories in demonstrating competence because accreditation alone does not guarantee quality. Therefore, many laboratories in various countries may be at different stages of implementing PT. PT is also essential for clinical trials and studies since data from variably competent laboratories can be misleading and sometimes is rejected by journals as an acceptable post for publication. PT is the preferable alternative to using samples from the general population or samples that have been artificially manipulated to contain abnormal results [1-3,5].

Proficiency testing (PT) is a system used to evaluate individual laboratories' work quality for specific tests or measurements and is widely used in clinical laboratories. It can be considered a valuable tool to reinforce quality assurance and cost-effective quality improvement. Any PT program compares a laboratory's performance with similar laboratories, regionally or nationally, and provides an educational experience to improve that performance. PT should not only be used to monitor ongoing competence but also to identify specific areas in need of further improvement [1,3].

Poor Performance

The definitions used vary; for example, Power (1996) defines a poor performer as a laboratory that fails to produce a result within defined standards. He recommended a percentage of satisfactory results using a z score (between -2 and 2) of 65% and 80%. Failure to achieve this was termed insufficient reliability. Blanket al. (2002) defined it as a result significantly different from the peer group consensus value. This definition has logical appeal and is based upon the premise in the Analytical Hierarchy that certain tests are so vital that the wrong result is worse than no result. Other researchers have measured performance in terms of categorical classification of results (Stableford et al., 2000) and diagnostic sensitivity (Shrader et al., 2004). Much research is not directly comparable because PP was studied concerning one method or disease-specific testing. PP is relevant to the on-site analytical process and results production; pre- and post-analytical influences may completely mask an error in the analytic phase. 6 This definition is good in that it highlights the fact that

poor performance affects patient results and can also be epitomized by instances where no result is worse than an incorrect one. The extent to which peer group performance is attained or error in a single test can be identified has informed the PT pilot study reported in this thesis and thus is closely related to the definition. The most widely used definitions are those of ISO 13528 (2005), the international standard for PT, which gives a quantitative definition of test-level bias and imprecision. An impressive performance of a method can be indicated by attaining the target clinical decision limit or complying with performance-based guidelines. ISO 13528 focuses on test-specific performance rather than results attained in routine analysis of patient samples. Thus, it may not accurately assess the influence of a method on patient results [6,7].

Significance of Investigating Poor Performance

Suppose these investigations of PT samples are to be effective. In that case, it will be necessary to know what evidence is that a particular result in a PT sample is incorrect. Usually, this will be implicit from how the sample has been 'solved' at the organizing laboratory; a discrepant result will be due to a certified difference between the target value and the result. 6Evidence can also be sought using the sample plus more general data from the scheme by trying to relate the performance of the sample to that in previous or subsequent schemes or to some related overall performance assessment in the laboratory, having first established what this represents. The ability to set up a PT sample investigation like any other analytical problem, varying the information available until the cost-effectiveness balance is right, is ideal. Yet, it is seldom possible to achieve this due to inadequate sample information or resources for a full investigation. However, an awareness of what evidence is available and the cost-effectiveness balance is necessary to avoid squandering resources on investigations that are doomed to failure [6].

Considerable thought should be given to investigating the causes of poor performance in proficiency testing. Such investigations consume considerable time and effort and add to the cost of external quality assessment - both for the laboratories concerned and for the organizers of the schemes. In one sense, it is a luxury to probe the errors identified in PT because, if resources were no problem, it would always be better to investigate and correct the causes of errors found in patient samples. However, as resources are finite, it

is necessary to take a pragmatic view and use them where they can do the best. PT samples are artificial, and the clinical risk to the patient due to their error is either nil or minimal. The same (or very similar) samples are often used in EQA schemes. So, repeating the exercise for the same sample type to investigate and correct an error found in PT may not be the best use of resources. However, we argue that it is usually worthwhile to investigate PT samples that have led to a misreport because they reflect an error that exists or could exist in patient samples and harm patients [8-10].

Factors Contributing to Poor Performance

Analytical errors have been cited as one of the most common reasons PT samples are reported incorrectly. Sub-optimal analytical performance is usually due to a failure to calibrate or standardize the method. This may be due to a lack of understanding of the importance of these processes. Often, there is no immediate feedback that errors have occurred, which can give a false impression of method performance. Failure to use the calibration materials provided by the manufacturer can also lead to errors in PT sample analysis. PT samples will often have different analyte concentrations than those of patient samples [8-10]. The use of the patient's internal quality control ranges to determine the acceptability of results in PT is a common cause of failure when the two do not correlate. Participation in PT, of which the results are known, can lead to a bias in methodology for PT samples to gain a desirable result. This can occur through deliberate and conscious manipulation of PT samples or by implementing costly but unnecessary changes to data management and reporting systems. PT samples are often treated differently than patient samples due to the urgency of reporting results or time spent engaging in additional problem-solving and consultation to resolve unacceptable results. Failure to treat PT samples as routine patient material and failure to simulate real-time analysis procedures is often the cause of unsatisfactory performance [9-11]. PT analysis results in feedback, and implementing corrective actions are important when errors in PT results occur. Often, these processes are not implemented correctly or at all. In rare cases, poor PT performance has led to the withdrawal of a laboratory's testing service to avoid scrutiny from regulatory bodies and/or litigation from patients [9,11,12].

Analytical Errors

Analytical errors can exhibit highly variable effects in a scale of PT/CAP survey samples, both random and systematic. Random errors cause the data to be scattered about an average value. Range statistics, standard deviation, and coefficient of variation measure the scatter of the data and are used to describe the precision of a method. Random errors are reflected in increased imprecision and decreased accuracy in the measurement process. This is seen when a method that normally achieves satisfactory results begins to perform poorly. A cause of the random error is any event or condition that follows the sample, which results in the taking of an incorrect measurement on that sample (i.e., spuriously high potassium values due to hemolysis of a serum sample in transit resulting from a poorly packaged shipment).^(8,10) In this case, the error occurred from failure to monitor an effective variable. Calibration errors are also classified as random errors, affecting the precision but not the accuracy of a measurement. If a systematic error is constant in both instances, it is possible to identify its presence when the laboratory plots its results on a Youdens Plot. It will be indicated by a result that falls significantly higher or lower than the true value. A change in the bias of the measurement will be detected by comparison with previous survey performance. This situation was observed in a clinical trial of patients presenting with septic symptoms. Despite no empirical evidence, the treating teams were convinced that the CRP of a patient should be vastly elevated in typical cases. The laboratory, therefore, increased the CRP reference range, and the data was compared with the previous cohort. Any significant change will be ultimately confirmed as an event of systematic error [10,12,13].

Instrument Calibration

While implementation of such recommendations on a widespread and sustainable basis will require significant resources and depend upon the availability of suitable PT materials, it is a neglected area that offers potential for improving the accuracy of results in some specific analytes. In summary, instrument calibration and monitoring for bias are key factors affecting performance in a wide range of assays, and laboratories need to consider using PT-based approaches to assess their current practice in this area. PT organizers must work with clinical scientists and regulatory bodies to provide appropriate support and guidance. Failure of a laboratory to achieve acceptable performance may well indicate the need to optimize the calibration of a method rather than any

shortcomings in the day-to-day process of sample and assay [14-17]. Recommendations have been made for using PT materials to monitor instrument bias in specific assays, particularly when there is evidence of method differences in specimen types and populations. This is a complex issue, particularly when disease states may be defined by specific assay results and clinical decisions based on these results. PT materials targeted at monitoring bias must be commutable for the samples likely to be encountered in routine practice, easily measurable by various methods, and reflect the clinical measurement interval for the assay.⁽¹⁸⁾ The importance of instrument calibration has been amply demonstrated in clinical chemistry methods subjected to standardization procedures. This was seen in the early days of cholesterol measurements when calibration of methods to reference methods resulted in very significant changes in reported values for the same specimens. More recently, changes in HbA1c measurement units to the IFCC scale have caused comparable shifts in reported values from the same specimens analyzed in different laboratories using different methods and have demonstrated the importance of comparability and the accuracy of results [19-21].

For the results to be accurate, laboratory instruments must be precise. Specially designed PT materials to assess instrument bias are not common in current PT schemes. However, while precision is a fundamental feature of all laboratory procedures and indicates a method's capability, proficiency testing measures the total effect, including both precision and bias [22]. Reagents used in the clinical chemistry laboratory must be of high quality. Because PT materials are often assayed by a method that includes comparison to a reference method (e.g., HPLC, GCMS, ICP-MS, IDMS) or a higher order material (e.g., atomic absorption, titrimetry), any bias in the value assigned to the PT material will impact upon the laboratory's analyte result. One source of error or bias potentially leading to inaccurate assignment of analyte values to PT materials is inappropriate or deteriorated reagents. This error may occur irrespective of whether the laboratory result is traceable to the PT material due to the state of the art in the assay; with the assignment of target values by peer groups for the commutable materials, the importance of this issue is increased [23].

Sample Handling and Storage

Sample mishandling and improper storage were often quoted as problems faced in the pre- and post-analytical phase and when incorrect results were reported. This study identified sample handling and storage as problems relating to poor PT performance [84-86]. In one instance, water contamination of a sample was traced back to a hole in the laboratory's roof. Unfortunately, the laboratory staff were unaware of the problem or the extent of the water damage to the sample until the EQA provider informed them of the outlier result [24, 25].

Operator Competence

The effect of poor operator technique and lack of experience on PT performance has been studied by collecting information from participating laboratories. Evidence of the Directorate General XII of the European Communities programs shows that high SD between results submitted by participants and the consensus mean result for a plasma pool generally agreed as outlier performance is associated with the employment of untrained staff in the analysis of CPD specimens [26,27]. The main problems encountered, which are due to insufficient staff training and experience, include failure to remove high molecular weight substances from ultrafiltration membranes before re-equilibration, interchanging serum and urine specimens, not following the correct analytical procedure, e.g., incorrect reagent dilution, and inability to make a clinical decision on whether a result is likely to be incorrect due to analytical error. Similar problems arose in the Scandinavian survey, where poor performance was significantly associated with staff having less than 2 years of experience in 6 of 13 subspecialties in medical biochemistry [15,28,29-31].

Mistakes made by entry-level personnel, which can drastically affect test results, are more likely to occur when staffing problems occur. In the USA, a survey conducted by Mallows et al. found that financial constraints imposed by managed care had forced many laboratories to lay off qualified personnel and fill vacancies with technical and clerical staff with little or no laboratory work experience. The evidence suggests that this shift in staffing levels may have led to increased errors and adverse events, detrimentally affecting patient safety. This is a worrying trend and one that has been recognized internationally. In response, the UK Council for Professions Supplementary to Medicine has set out a policy for the training and development of HCS staff, which should be closely adhered to if patient safety is not to

be compromised. The influence of staffing levels in laboratory error is of further concern for CPD analysis, as some laboratories may use a general clinical chemist, who may leave the post because of retirement or career advancement, resulting in a replacement with a specialist in a different area of medical biochemistry. A decrease in error may be seen if medical biochemists reporting poor performance in subspecialties outside their area of training change to the analysis of tests in their field of knowledge [32,33].

Pre-analytical Errors

The most common errors that cause poor PT performance are misidentification of patient samples, deteriorated or contaminated samples, inappropriate sample collection, unsuitable test requests, or incorrect test selection. Sample mix-ups are common within clinical laboratories and may go undetected unless an unexpected PT result is obtained. It can occur within or outside the laboratory when samples are taken to the laboratory and left in a location where other patient samples are also present, for example, a ward or a doctor's surgery. The effect of a mixed-up sample is that the laboratory result will not be for the true patient and can cause downstream negative effects for the patient whose sample was incorrectly analyzed. In some cases, this type of error may be discovered if patients' samples are labeled in any way that allows the sample identity to be electronically verified and tracked through any automated analyzer [34-36].

Special attention should be given to the difference between pre-analytical errors and the other types of errors, which are problems that occur before the diagnostic analysis. On the other hand, analytical errors occur during testing, and post-analytic errors occur after testing. When investigating poor PT performance, it is only possible to determine that an error is likely to have occurred. It is not always possible to definitively diagnose the exact problem. Several systematic techniques can be used to identify the nature of specific errors. These include reviewing the entire testing process, seeking any anomalies, parallel testing, blind split testing, and standard addition of known analytes at different concentrations to patient samples. These methods involve re-testing patient samples while manipulating test conditions and comparing results to determine where a problem may have occurred. ^(6,34,35,37)

Post-analytical Errors

The Dempster review reported that deviations from preceding qualitative elements could lead to post-analytical errors that contribute to diagnostic errors. One of the main functions of the clinical biochemistry laboratory is to provide accurate and reliable data, and this is achieved through the utilization of biological variation data and other available evidence to set analytical quality specifications, which are then used to measure the performance of an analyte or measurement procedure. This can be assessed during internal quality control, during which the laboratory will measure the precision of the data gathered. Here, they use the guidelines provided by biological variation data to decide whether the precision of an analyte is acceptable, with limits set to designate a desirable or minimum standard [37, 38]. Although imprecision is the main focus of biological variation when assessing analytical performance, accuracy is also assessed through bias, and this is a key area where post-analytical errors can be identified. External quality assessment is a procedure used in the UK and other countries worldwide to assess and improve the performance of a laboratory in the analytical phase by having a second party assess the quality of the results of tests made by the first party. This is largely based on proficiency testing, where a national scheme provider sends simulated patient samples to the laboratory, which will then peer-review the laboratory's performance [34,36-38]. When the laboratory investigates the poor performance of an analyte in a proficiency test, one process that differentiates between analytical and post-analytical errors is to compare the analyte's performance with the biological variation data or a set of predefined analytical quality specifications. This process may reveal that an acceptable level of precision and trueness has been achieved, demonstrating that the cause of the poor performance is an error in the interpretation of the clinical information given, or it may show that the analyte has deviated from the quality specifications indicating an analytical error [34,39,40]. Other methods to distinguish between the two error types include a review of the analytical phase and how the data was obtained, and investigations often use a fishbone diagram to map out the various influences on the production of analyte data, as seen in a study by Carlos Rios. The understanding of the error type is crucial as a post-analytical error may have serious implications on the patient if the clinical decision is influenced by inaccurate or misinterpreted

data, and it is, therefore, necessary to act on the error and attempt to improve the performance of the analyte. This is evident by the recent decision by the Australian government and the Institute of Clinical Pathology and Medical Research to underpin all of the pathology industry in Australia with a quality framework based on Risk Management and the recent publication of an ISO standard for medical laboratories based around the same thing [34,39,41]. Sample mishandling and improper storage were often quoted as problems faced in the pre- and post-analytical phase and when incorrect results were reported. This study identified sample handling and storage as problems relating to poor PT performance [24,25,26]. In one instance, water contamination of a sample was traced back to a hole in the laboratory's roof.⁽²⁴⁾ Unfortunately, the laboratory staff were unaware of the problem or the extent of the water damage to the sample until the EQA provider informed them of the outlier result [24-26].

Staff Competency and Training

Staff inexperience and lack of knowledge are significant issues leading to incorrect method selection or inappropriate test performance, resulting in calibration, control, and patient sample results differing from the true value. One way to combat this is through appropriate education and training. Staff should regularly attend appropriate training courses and be encouraged by laboratory management to strive for further education in clinical chemistry. Continuous education will improve staff competency, but proficiency testing serves as continuous education and competency improvement. Regular participation in an external PT program will force staff to identify and resolve issues in their analysis system, thus using PT analytes as a "practice" for testing new methods or troubleshooting existing ones. PT will also help determine whether a new staff member can perform testing independently [34, 42,43]. If a laboratory has been notified of unacceptable performance in PT, corrective action should be taken, and staff competency should be reassessed. It is necessary to establish performance criteria and set limits to determine whether a result is acceptable. This is often done by comparing a known true value to the laboratory result, effectively the same as PT using a patient sample. If the performance is still unsatisfactory, then that particular testing should be left until further training or verification of the method can be performed. In some cases, it may be necessary to replace staff [34,40,43]. Overall, the most

important factor in staff competency and training is the attitude, common goal, and response to errors in the laboratory. This requires full cooperation between lab management, staff, and pathologists. Errors in analysis or PT must be met with emphasis on the need for error identification and resolution and not punishment. The laboratory management should determine the cause of the error and what action can be taken to prevent a recurrence and then catalyze the changes. PT provider contact can also be helpful, providing a resource to determine the correct result and the most appropriate method to obtain it [34,39,42].

Methods of Identifying Poor Performance

Internal quality control (IQC) in the laboratory can be an effective tool in identifying poor laboratory performance in between PT exercises. It can be more effective than comparing results directly to the assigned value as it will be more sensitive in identifying problems. IQC is defined as the set of procedures used to monitor the analytical process after it has been established that the process is adequate to maintain the quality of results. This is an ongoing method of PT monitoring, which can be just as effective in identifying patient problems and PT results. IQC methods can include setting control limits for the target of measurement, producing a control chart to monitor changes in the accuracy and precision of the target, or it can involve comparing an internal control material to another method of measurement to validate the method in addition to monitoring its long-term bias and imprecision [34,41,43]. Statistical analysis of proficiency testing (PT) results is recommended by the International Organization for Standardization (ISO) to identify poor laboratory performance. It is a simple and objective method of identifying outliers and trends in PT results. It has the advantage that the performance of the laboratory can be compared to an objective target, i.e., the assigned value, Δ . Various statistical methods have been employed to identify outlying PT results, such as modified z-scores and the Levey-Jennings quality control procedure. Z-scores are the most commonly used statistical method, with very few laboratories using other tests that can be just as effective, if not more so. There are many methods of calculating z-scores, many of which are inappropriate for PT results

Statistical Analysis of Proficiency Testing Results [8,10,40,44]

The error types analysis is still in development for the strategy to identify poor performance. The error can be defined as deviations from the true value caused by bias, imprecision, and total (random) analytical errors that decrease the probability that the result will be taken as no clinical change when it indicates significant clinical change. This condition usually leads to a failed or wrong decision in interpreting the patient's clinical data or medical decision, but it could be an unseen error. The step to identify the error type is to compare the laboratory results with the PT provider's and the other peer group's results. This analysis will need vast information from the PT provider and other peer groups. By comparing those results, a pattern of error types could be identified much better [45].

This is the easiest, lucid, and efficient way of identifying poor performance compared to peer groups. It is explained by mining various statistical information from the PT result, such as z-score, rate of correction, and error types for every analyte. The z-score is usually calculated using this equation: $Z = (x - X) / s$, where x is the laboratory result, X is the mean result, and s is the standard deviation. The z-score will measure how many standard deviations between the laboratory result and the mean result. The rate of correction could be calculated using a percentage. This calculation requires a specific value of the decision about the acceptability or non-acceptability of the rate. The rate of correct is the proportion between the number of proficiency results classified as correct divided by the number of proficiency testing events [46,47,48].

External Quality Assessment Programs

The result of the quality proficiency test from the participant, as seen from the graphic and table, doesn't show the real condition of the test. But, with the previous information, it can give suggestions and feedback towards the test so the next PT can be done with better preparation. The other statistical data processing can make a table comparing the glucose test from the participant with other participants and a histogram comparison of both tests. It's based on the graphical and table information about test quality from the participant, which can be easily approached [49,50].

With the data, a Clarke error grid consisting of the regression line of standard criteria vs participant results can be made. The picture below describes the glucose result of the participant compared with standard criteria and the obtained regression line. We

obtain information about bias or systematic error from the intercept of the regression line, which is -73.105. If stated as LTC value, the glucose test has BTC since the regression line is nearing the continuum of standard criteria. Then, we calculate the correlation value with the real state correlation of 0.913 and SD of the regression line 7.402. These two values can be used to calculate the value of %TE from the formula $125\% \sqrt{(1+r)(1-s)}$, whereas $s = \sqrt{(n-2)/n} \sqrt{1/n \sum (x_i - \bar{x})^2}$ for the obtained n value and x_i is the LTC value. After getting the value of TE current and slope from the regression line previous Std criteria-LTC, %TE uncertainty plot vs BTC can be made, which is the value of the slope line $\pm te/s \sqrt{1-r^2}$ with an ellipse shape and the location of that glucose test point. Lastly, %bias and %TE report of the glucose test complete with a histogram and findings from the beginning [49,51-52].

To mediate work for quality improvement, each participant should be identified and occasionally give immediate suggestions for their program. As described previously in linearity analysis, the presence module total error estimation system can potentially suggest PT performance based on the standard criteria value. During the following module, several graph participants taken from PT result 2005 of participants with code CMS 331 will be displayed. The use of MS Excel combined with the 2005 Excel add-in program from [Link] to process the data aimed to simplify the statistical calculation of each test from the participant. The Excel data table filled with the coding of standard criteria and LTC and %bias values from the participant, which was processed with error grid conversion, can be seen as an xy scatter plot of each test. We take the example of the glucose test. (54-57)

Internal Quality Control Measures

Quality control is part of the quality management process focused on fulfilling quality requirements. To guarantee the reliability of the laboratory results, those results must be precise (free from random error) and accurate (free from both systematic and random error). It is related to the testing process in that precise and accurate results are based on properly functioning reagents and instruments and on technicians who follow correct procedures. Errors in test results can occur due to faulty procedures, inexperience of the technician, or improperly functioning equipment [59-61]. To find and correct errors as they occur, the clinical chemist needs reliable and rapid methods to assess the precision and accuracy of the testing process. Control materials are used to monitor

precision and accuracy on a day-to-day basis. Precision is assessed using control charts, where it is expected that 1/3 of results will fall outside the range $\pm 2s$, and 1/20 will fall outside the range $\pm 3s$ (where s is the standard deviation of the results). A large group of control results should form a normal distribution. Multiple control materials at different concentration ranges should be tested. Accuracy is assessed by comparing the laboratory's results to the true value of the control material. If the control material is human blood or urine, it may enrich known samples, and an assay on patient samples can double as an accuracy assessment [62-66]. Any detected imprecision or inaccuracy must be promptly dealt with. Reagents, equipment, and technician methods must be evaluated and corrected, and if errors persist, new reference ranges may be established. Control material analysis is a form of internal quality assessment, and to ensure that the overall quality of testing is acceptable, the clinical chemist should participate in external quality assessment programs [65,67,68].

Review of Control Charts

The initial step in any investigation involves identifying the issue, such as discovering any abnormal patterns or changes in the test performance. The laboratory conducts a wide variety of tests, and to enhance efficiency and speed, data from multiple analytes may be examined simultaneously. However, for the EAP, each analyte should be evaluated individually. Established control charts are the most effective method for identifying significant alterations in test performance. Nelson's rules have been devised for numerous benchmarked EQUAS schemes, such as the Ricos trueness verification protocol, which considers biased and total allowable errors to measure systematic error. An online tool called the quality and error control expert (QUALEX) has been specifically developed to identify analytical errors in EQUAS data. QUALEX offers an advantage over other Shewhart chart rule systems because it allows for easy inclusion or modification of rules to investigate specific errors derived from EQUAS data results [69-72].

This sets out the guidelines to be followed by the laboratory whenever EQUAS-like data is analyzed. The EQUAS scheme aims to monitor the results of routine clinical laboratory testing. Consequently, to truly reflect routine performance, all available data and control results should be included in the analysis. This inclusiveness is a fundamental difference between an evaluation and an investigation. The

more failures there are to include relevant information, the more the investigation becomes an evaluation of the failure itself. Any investigation must find problems that can be corrected rather than just identifying failure. The findings should be of value to other laboratories that may be experiencing similar problems, even though the specifics of the problems may differ. The quality coordinator will conduct All data analysis and may delegate specific tasks to other staff with suitable expertise. [73-75].

External Quality Assessment Programs

Probability-based models and multilevel regression are used to evaluate the EQA and the potential effects of covariates. This new evaluation method eliminates further costs and time and can help assess systematic and random errors in data generated from PT events. This more recent approach to post-analytical PT data assessment has provided the opportunity to fully understand errors that occur during PT events and assess the impact and outcome of corrective action [16,76-79]. Adverse analytical events (AAE) have been defined as medical laboratory errors detected before, during, or after using laboratory information to make decisions for patient care. This led to the development of a new method in which PT samples do not require formal testing to obtain results. Instead, they observe patient samples or retrospective data that have used the same tests, methods, and reagents on a routine basis. This will allow for the comparison and analysis of data collected from AAE with the PT event data [80-82]. The two most commonly applied methods of external quality assessment for PT programs involve the post-analytical approach and the data and their data evaluation. The post-analytical approach involves re-testing PT materials in routine diagnostic laboratory conditions and comparing these results with those obtained during the PT event. Any discrepancies in results can then be investigated, and corrective action can be taken. This approach effectively determines the cause of error but is expensive and time-consuming [83-86].

Peer Review and Inter-laboratory Comparisons

This method of identifying sub-standard performance is probably one of the best, but regrettably, it is rarely done. The idea is that data on methods and results of a particular analysis can be shared between two or more laboratories. This data can then be critically evaluated. In doing so, laboratories could identify and correct local problems in their analytical process. The

main criticism of this method is that it has proved too costly for many laboratories to undertake and has a poor turnaround time. Since PT has become increasingly easier and cheaper to access, it may be feasible to begin implementing PT comparisons as a mandatory part of a PT program. An example of how successful an inter-laboratory comparison can be as a tool of improvement can be seen in the Swiss Vitamin C PT program. This program included a matrix of 4 different lyophilized sample types, each distributed to the participating laboratories. An evaluation of the results from this trial showed significant improvement in performance regarding accuracy towards the consensus mean [87,88,89]. As we can see, there are several ways that laboratories' performance can be identified and improved upon. PT results provide an undeniably good data source that will allow tailored guidance to laboratories, as it is often difficult to see the areas of weakness in the analytical process. The main problem at this point is associating poor performance with the cause of error and rectifying the problem, which in many cases will be a lack of motivation due to little consequence of failure. If laboratories are unable to apply measurable suggestions and witness improvement, there is little chance their performance will change [90].

Strategies for Addressing Poor Performance in Laboratory Testing

A strategy that seems to be often overlooked is the progression from addressing poor performance to continuous improvement. Most case studies look at how to return to an acceptable level of performance. This is usually due to a lack of perceived benefit from comparing the effectiveness of different strategies against each other and the cost of implementation. One case study demonstrating the opposite approach is investigating poor performance in the pilot scheme of a new EQA program for cystic fibrosis genetic analysis. This was to replace their current method, and it was decided that there would be a greater long-term benefit to investigating and addressing the poor performance as it would affect the laboratory's ability to offer reliable testing for cystic fibrosis patients. The expectation was to identify the cause of failure and an opportunity to compare different strategies for successfully implementing the new method [91,92]. Considering the occurrence of poor performance in subsequent PT schemes, an important consideration for an institution is using resources to investigate and address the poor performance. This may involve cessation of routine work to free staff and time to

investigate the cause of poor performance and satisfactory completion of the investigation to enable changes to be implemented. This may have cost implications if performance in PT is linked to financial incentives or reimbursement rates, and loss of revenue triggers a sequence of events, eventually costing more than the original lost revenue. An approach such as this was taken by the Chemistry Laboratory at Emory Healthcare and can be related to their decision to withdraw from accreditation and reapply later. Changes implemented due to the investigation must effectively ensure errors are not repeated. This may involve a trial-and-error approach to find the most appropriate corrective action. An example given in the case studies is the change from reagent rental to purchase to improve the calibration and maintenance of equipment. [93,94].

Thus, readers should understand why laboratories show poor performance in PT, what the procedures involved, and who is responsible for outlining strategies to address poor performance. They should also understand how some of the case studies presented in other sections could have been approached differently. [94,95]

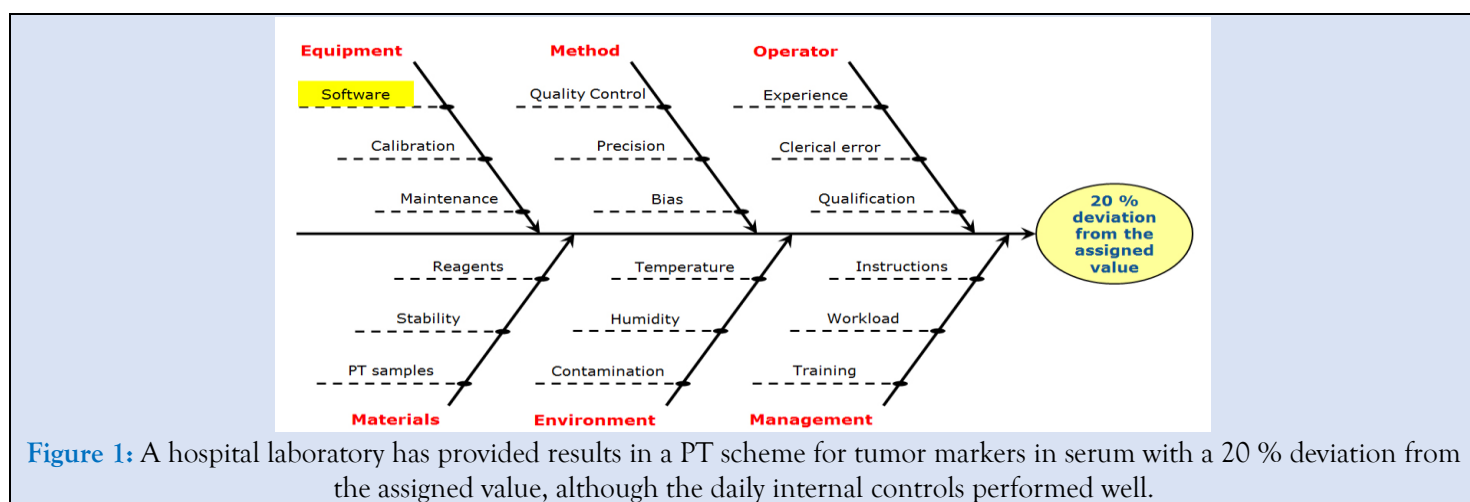
Root Cause Analysis

An RCA model looks for underlying reasons or "root causes." Thus, it extends beyond pinpointing the root causes of performance, quality, or safety issues. The fundamental cause of an issue that has arisen or a deviation or non-conformance that has been found has been identified as the root cause. A variation may have one or more underlying causes. To improve systems or processes, one has to understand the underlying cause. If the root cause of the issue is not found and fixed, it is likely to recur or, at most, affect other parts of the system or process. The goal of traditional linear root cause analysis [RCA] models has been to identify one or a few fundamental causes [94,95]. The significance of the analysis, the frequency of disappointing results, and the presence of trends will all influence how in-depth the investigation becomes. The lab ought to verify if the PT report explains the subpar performance. A progressive approach is recommended to increase the likelihood of identifying the problem's core cause in the absence of a justification [96,97]. The staff who performed the analysis, as well as, if applicable, the management of the laboratory, ought to be involved with the following steps of the inquiry: Analyze the raw data, data from internal quality control, any patterns from earlier PT rounds, and the participants' overall

performance in that round; Create a plan for corrective measures after the inquiry is finished, taking into account any implications for earlier test results; carry out and document the remedial measure or actions; Verify the efficacy of the corrective remedy or actions [95,96].

The staff who carried out the analysis as well as, if applicable, the management of the laboratory, should be involved in the following steps of the investigation: Examine the raw data, data from internal quality control, any patterns from earlier PT rounds, and the participants' overall performance in that round; Create a plan for remedial action after the

investigation is finished, taking into account any implications for earlier test results; take out and document the remedial measure or actions; Verify the efficacy of the corrective solution or actions. Problems related to the PT scheme. Poor performance could also be due to the PT scheme not being entirely appropriate. The Eurachem Guide provides information on selecting an appropriate PT scheme. In other cases, a problem may have occurred with the PT items. The laboratory is encouraged to discuss its findings with the PT provider [95,96]. The “Fishbone diagram” or “5 Whys” are useful tools to investigate the root cause of a problem Figure 1,2.



The laboratory investigated the process and discovered the mistake was due to an incorrect unit conversion factor introduced for PT results when the IT system was last updated. Since patient results are reported in a different unit than the one used by the PT provider, no patient results were affected [95,96]. *A laboratory performing microbiology analyses obtained poor performance scores because the number of microorganisms found was underestimated. Laboratory investigations showed an issue with the incubation time and temperature applied by the analyst. The laboratory found that the analyst used an out-of-date procedure [95,96].*

1. Poor growth of *Listeria monocytogenes* **why?**
2. Incorrect incubation time and temperature **why?**
3. The applied test procedure contained a mistake ... **why?**
4. An outdated version of the procedure was used .. **why?**
5. The revised procedure was not distributed **why?**

The workload of the person in charge of document control

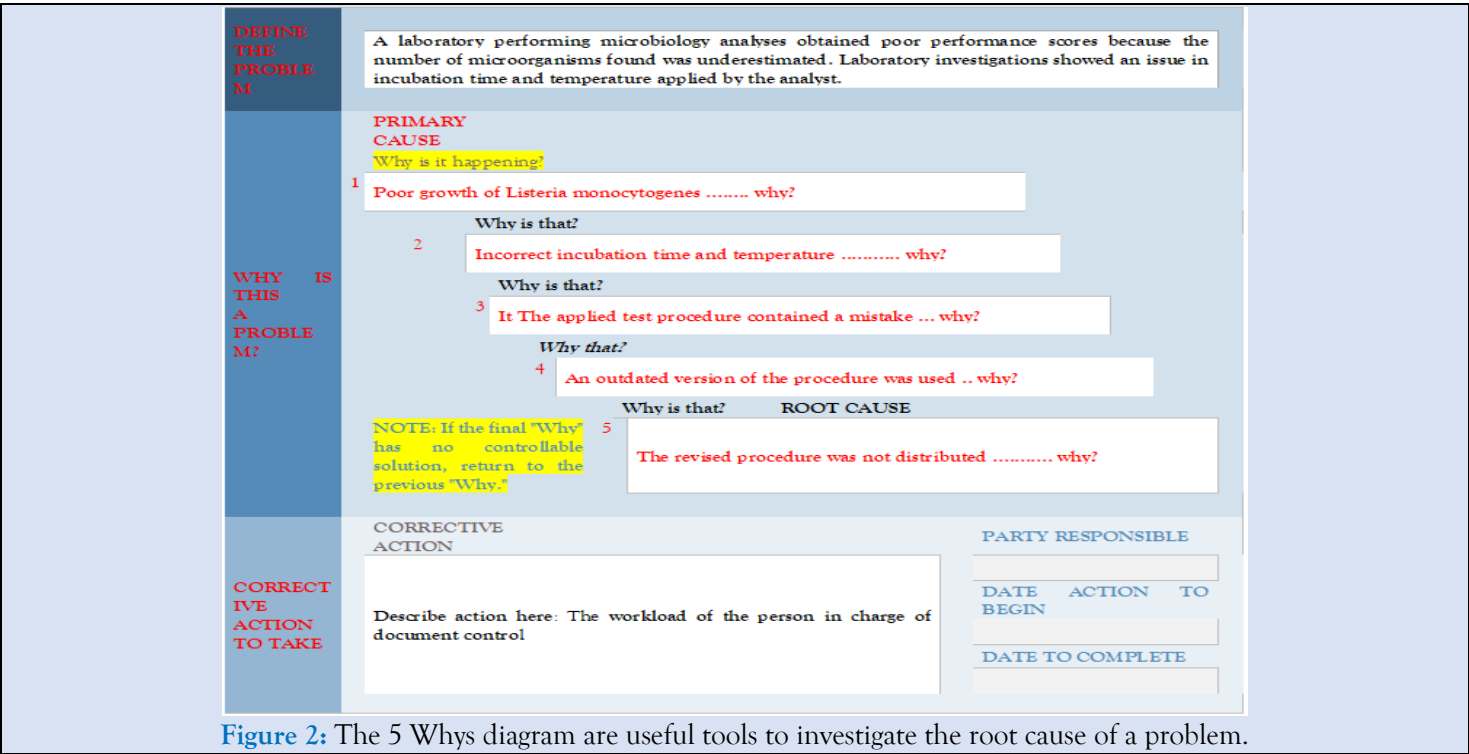


Figure 2: The 5 Whys diagram are useful tools to investigate the root cause of a problem.

Corrective and Preventive Actions

Steps of corrective and preventive action: (i) Define the event Make an objective for the event to identify the problem and measure the effectiveness of improvement at the end of preventing the event. (ii) Assemble a team. Select an active team or assign responsible personnel to identify the cause of an event. The team must work together in root cause analysis to ensure a systematic understanding of another event's cause and effect. (iii) Find out the root cause. Use tools in RCA, such as a fishbone diagram and 5 why, to assist in understanding the event cause. This will enable easier planning of corrective and preventive action [97-99]. (iv) Determine the corrective action. Ensure the action implemented only rectifies the event's cause and does not change any other part that might affect the rectified event. A good plan will be specific, realistic, with a timeline,

systematic, and measurable; it involves basic principles such as "What will be done?" and "How is it going to be achieved?" Corrective action is an action taken to rectify and eliminate the cause of actual poor performance in a running event to prevent the occurrence of such an event in the future. In contrast, preventive action is an action to avoid such an event [98-100]. Several measures can be taken to address poor performance to run preventive and root cause action, such as removing the personnel, modifying the method and routine, purchasing and installing a new instrument, and providing additional training and education. All of the measures should be to increase the testing accuracy of the event. Preventive action should be created as a policy that enables the improvement taken to be long-lasting and not just solve the event only. These actions must be monitored, and their effectiveness should be assessed simultaneously [100,101].

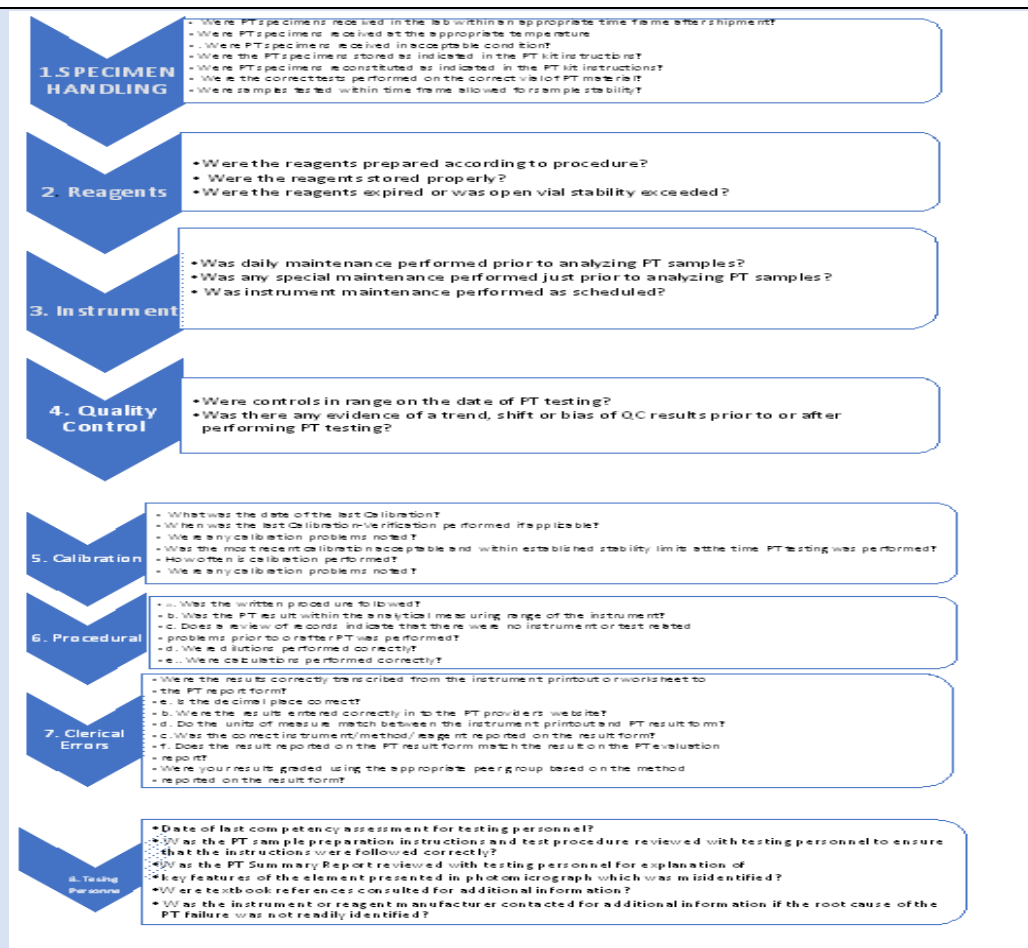


Figure 3: Check Steps for Proficiency Testing Corrective Action

Prenalytical Errors: A "No" response may indicate specimen handling issues, often due to failure to read PT kit instructions.

Analytical/Procedural Errors: A "No" response may suggest analytical or procedural errors, possibly due to not performing recommended instrument maintenance/calibration or not following written test procedures. Review PT kit instructions or standard operating procedures as needed.

Clerical Errors: A "No" response may point to clerical mistakes. Though different from patient result reporting, these errors may require additional staff training, instruction review, adding a second reviewer, or investigating the reporting format. Contact your proficiency testing provider if the reported results do not match the evaluation report.

Continuous Improvement Initiatives

The choice of which projects to undertake is important, and FOCUS-PDCA is a decision-making tool that can be used in selecting and managing improvement projects. This structured problem-solving methodology splits the process into two main phases and has a systematic review built in at each stage to judge whether the continued effort is

worthwhile. First, "Focus" involves defining the problem and setting an objective that must be achievable, and then "PDCA," which stands for "Plan-Do-Check-Act," is the process for reaching the objective. At each instance of "Act," the results should be reviewed to determine whether to proceed to the next step [101-103]. By conducting a "Check" at the end of the first phase, it can be determined whether the issue is within the realm of influence of the laboratory and if the problem definition and objective are correct. This is important as it avoids wasted effort, and the "Check" or review is a recurring theme similar to establishing the cause of issues in the CAPA process. By performing continuous improvement initiatives, general problem-solving will resume, and the findings from RCA and CAPA can be implemented into changes in processes and monitored for their effectiveness [101,103]. Incorporating continuous improvement into the culture of the laboratory is essential for ensuring that corrective actions are appropriate and effective and that performance is maintained or enhanced in the long term. However, to be effective, a structured program of projects or initiatives driven by

information from the past aimed at improving processes should be established. This is in contrast to a strategy of relying on "firefighting," in which action is taken only in response to an issue and frequently results in ill-directed effort. An example of this is where a sudden peak in failure of a particular analyte triggers a flurry of activity directed at troubleshooting the analytical system for that analyte. Yet, the cause of the problem may be found in the pre-analytical phase or even in how the EQA samples are received and managed [102,103].

Staff Education and Training

One of the most significant factors in causing errors is the relative lack of training given to laboratory personnel. This contrasts the complexity of the tasks they are expected to perform. In the RT scheme, assessments are made of laboratory staff competency and its relation to laboratory performance. The methodology used in this has been to measure the cognitive skills of the laboratory staff, i.e., the ability to understand and answer questions on the theory underpinning a given clinical biochemistry assay. In the UK, laboratory employees are now expected to know the underlying principles of clinical biochemistry [102-105]. This follows the successful implementation of the ACB National Vocational Qualifications [NVQ] scheme, a well-structured, certificated form of training with defined objectives recently incorporated into the IBMS training program. Third, the increasing dependence of laboratories on automation will help reduce human error. Still, it raises machine-related mistakes, necessitating a higher level of staff expertise to diagnose and remedy these when they occur. After assessment of the training and educational needs of laboratory personnel, it is recommended that public awareness of the importance and complexity of laboratory medicine be increased to encourage the provision of adequate training resources. Steps should be taken to assign responsibility for RT and implement its findings. An RT coordinator could be appointed for a given clinical area, e.g., diabetes, to monitor changes in performance, maintain contact with PT providers, and investigate any variability in laboratory performance seen between different types of samples. This work cannot take place without the necessary funding, so laboratory managers and clinicians must be made aware of the benefits and cost-effectiveness of RT through the provision of hard evidence as to its efficacy [5,104,105].

Conclusion

In conclusion, this review underscores the importance of investigating and addressing unacceptable results in proficiency testing among medical laboratories. Such results indicate potential issues in laboratory performance, with implications for patient care and safety. To mitigate these challenges, laboratories must implement robust quality assurance measures, including staff training, adherence to standardized procedures, and regular equipment maintenance. Participation in proficiency testing programs is also vital, providing valuable insights for continuous improvement. Collaboration among laboratories and ongoing monitoring is essential for identifying and addressing common challenges, ultimately enhancing the quality and reliability of laboratory test results for improved patient outcomes.

References

1. Tziakou, E., Fragkaki, A. G., & Platis, A. (2023). Identifying risk management challenges in laboratories. *Accreditation and Quality Assurance*, 28(4):167-179.
2. Abolhasani, M., & Kumacheva, E. (2023). The rise of self-driving labs in chemical and materials sciences. *Nature Synthesis*, 2:483-492.
3. Miller, W. G., et al. (2023). Overcoming challenges regarding reference materials and regulations that influence global standardization of medical laboratory testing results. *Clinical Chemistry and Laboratory Medicine (CCLM)*, 61(1):48-54.
4. Taylor, C. J., et al. (2023, February 23). A brief introduction to chemical reaction optimization. *Chemical Reviews*, 123(6):3089-3126.
5. Atkins, D., et al. (2022, May). Accelerating battery characterization using neutron and synchrotron techniques: Toward a multi-modal and multi-scale standardized experimental workflow. *Advanced Energy Materials*, 12(17):2102694.
6. Hawk, J., Ray, S. J., Robinson, J. K., & Thomas-Smith, T. E. (2022). Multi-week lab projects incorporate breadth and depth into analytical chemistry laboratory courses. *Inactive Learning in the Analytical Chemistry Curriculum. American Chemical Society*, 221-248.
7. Rushing, G., Power, B., & D'Alessio, M. (2023). Implementing service-focused activities focused on water quality while teaching a freshmen

- undergraduate course during the COVID-19 pandemic. *Journal of Chemical Education*, 100(5):1843-1851.
8. Vounba, P., Loul, S., Tamadea, L. F., & Siawaya, J. F. (2022, March 31). Microbiology laboratories involved in disease and antimicrobial resistance surveillance: Strengths and challenges of the central African states. *African Journal of Laboratory Medicine*, 11(1):1570.
 9. Bai, J., Cao, L., Mosbach, S., Akroyd, J., Lapkin, A. A., & Kraft, M. (2022, January 10). From platform to knowledge graph: Evolution of laboratory automation. *JACS Au*, 2(2):292-309.
 10. Wenzel, T. (2023, January 17). Open hardware: From DIY trend to global transformation in access to laboratory equipment. *PLoS Biology*, 21(1):e3001931.
 11. Maphalala, M. C., Mkhasibe, R. G., & Mncube, D. W. (2021). Online learning as a catalyst for self-directed learning in universities during the COVID-19 pandemic. *Research in Social Sciences and Technology*, 6(2):233-248.
 12. Garces, D. G., et al. (2023). On the reproducibility of extrusion-based bioprinting: Round robin study on standardization in the field. *Biofabrication*, 16(1):015002.
 13. Flowers, R. M., et al. (2023). (U-Th)/He chronology: Part 1. Data, uncertainty, and reporting. *Bulletin*, 135(1-2):104-136.
 14. Panteghini, M. (2022). Redesigning the surveillance of in vitro diagnostic medical devices and of medical laboratory performance by quality control in the traceability era. *Clin Chem Lab Med*, 61(5):759-768.
 15. Teshome, M., Worede, A., & Asmelash, D. (2021). Total clinical chemistry laboratory errors and evaluation of the analytical quality control using sigma metric for routine clinical chemistry tests. *Journal of Multidisciplinary Healthcare*, 125-136.
 16. Abdel, G. M. T., & El-Masry, M. I. (2021). Verification of quantitative analytical methods in medical laboratories. *J Med Biochem*, 40(3):225-236.
 17. Badrick, T., Bietenbeck, A., Katayev, A., van Rossum, H. H., Loh, T. P., & Cervinski, M. A. (2020, November 16). Implementation of patient-based real-time quality control. *Critical Reviews in Clinical Laboratory Sciences*, 57(8):532-547.
 18. Luks, A. M., & Swenson, E. R. (2020). Pulse oximetry for monitoring patients with COVID-19 at home: Potential pitfalls and practical guidance. *Ann Am Thorac Soc*, 17(9):1040-1046.
 19. Cabitza, F., et al. (2021). The importance of being external: Methodological insights for the external validation of machine learning models in medicine. *Computer Methods and Programs in Biomedicine*, 208:106288.
 20. Bohn, M. K., Lippi, G., et al. (2020, June 25). Molecular, serological, and biochemical diagnosis and monitoring of COVID-19: IFCC task force evaluation of the latest evidence. *Clinical Chemistry and Laboratory Medicine (CCLM)*:58(7):1037-1052.
 21. Stevens, R. J., & Poppe, K. K. (2020). Validation of clinical prediction models: What does the calibration slope really measure? *J Clin Epidemiol*, 118:93-99.
 22. Pashmdarfard, M., & Azad, A. (2020). Assessment tools to evaluate activities of daily living (ADL) and instrumental activities of daily living (IADL) in older adults: A systematic review. *Medical Journal of the Islamic Republic of Iran*, 34, 33.
 23. Naphade, M., Ankur, P. S., & Rajendra, S. R. (2023). Quality control in clinical biochemistry laboratory—A glance. *Journal of Clinical & Diagnostic Research*, 17(2).
 24. Van Mullem, T., et al. (2020). Addressing the need for standardization of test methods for self-healing concrete: An inter-laboratory study on concrete with microcapsules. *Science and Technology of Advanced Materials*, 21(1):661-682.
 25. Ratnamurni, E. D., Ludiya, E., & Luthfiartie, A. (2022, January 20). Quality risk management in infusion product distribution using Failure Mode and Effect Analysis (FMEA) and Analytical Hierarchy Process (AHP) methods. *International Journal of Science, Technology & Management*, 3(1):229-244.
 26. Wan, X., Xie, G., et al. (2023). National proficiency test for molecular detection of white spot syndrome virus in China, 2021. *Aquaculture*, 573:739579.
 27. Raurell-Torredà, M., et al. (2021). Effectiveness of a training intervention to improve communication between/awareness of team roles: A randomized clinical trial. *Journal of Professional Nursing*, 37(2):479-487.

28. Brookman, B., et al. (2024). Proficiency testing in analytical chemistry, microbiology and laboratory medicine: Discussions on current practice and future directions. *Accreditation and Quality Assurance*:1-9.
29. Arifin, A., & Mohd-Yusof, M. (2022). Error evaluation in the laboratory testing process and laboratory information systems. *J Med Biochem*, 41(1):21-31.
30. Graden, K. C., Bennett, S. A., Delaney, S. R., Gill, H. E., & Willrich, M. A. (2021). A high-level overview of the regulations surrounding a clinical laboratory and upcoming regulatory challenges for laboratory-developed tests. *Laboratory Medicine*, 52(4):315-328.
31. Fenta, D. A., & Ali, M. M. (2020). Factors affecting quality of laboratory result during ordering, handling, and testing of the patient's specimen at Hawassa University College of Medicine and Health Science Comprehensive Specialized Hospital. *J Multidiscip Healthc*, 13:809-821.
32. Antonios, K., Croxatto, A., & Culbreath, K. (2021). Current state of laboratory automation in clinical microbiology laboratory. *Clin Chem*, 68(1):99-114.
33. Thakur, V., Akerele, O. A., & Randell, E. (2023). Lean, and Six Sigma as continuous quality improvement frameworks in the clinical diagnostic laboratory. *Critical Reviews in Clinical Laboratory Sciences*, 60(1):63-81.
34. Mirowsky, J. E. (2020). Converting an environmental sampling methods lecture/laboratory course into an inquiry-based laboratory experience during the transition to distance learning. *Journal of Chemical Education*, 9:2992-2995.
35. World Health Organization. (2020). Laboratory biosafety guidance related to coronavirus disease (COVID-19) outbreak. *World Health Organization*.
36. Koh, L., & Chia, Y. T. (2023). Automated cell culture monitoring and diagnostic assay: A systematic review. *Biotechnol Progress*, 40(4):e3279.
37. Brewster, D. T., et al. (2023). Safety and perceptions of risk in the handling of laboratory chemicals in a biological research community. *ACS Chemical Health & Safety*, 30(2):63-71.
38. Burstyn, I., Goldstein, N. D., & Gustafson, P. (2020). Towards reduction in bias in epidemic curves due to outcome misclassification through Bayesian analysis of time-series of laboratory test results: A case study of COVID-19 in Alberta, Canada and Philadelphia, USA. *BMC Medical Research Methodology*, 20:1-10.
39. Tran, K., Beshir, A., & Vaze, A. (2021). A tale of two lab courses: An account and reflection on the teaching challenges experienced by organic and analytical chemistry laboratories during the COVID-19 period. *Journal of Chemical Education*, 97:3079-3084.
40. Choate, J., et al. (2021). International educators' attitudes, experiences, and recommendations after an abrupt transition to remote physiology laboratories. *Advances in Physiology Education*, 45(2):310-321.
41. Sussman, E. M., et al. (2022). Chemical characterization and non-targeted analysis of medical device extracts: A review of current approaches, gaps, and emerging practices. *ACS Biomaterials Science & Engineering*, 8(3):939-963.
42. Rappold, B. A. (2022). Review of the use of liquid chromatography-tandem mass spectrometry in clinical laboratories: Part II-Operations. *Annals of Laboratory Medicine*, 42(5):531-557.
43. Poo, M. C. P., Lau, Y., & Chen, Q. (2023). Are virtual laboratories and remote laboratories enhancing the quality of sustainability education? *Education Sciences*, 13(11):1110.
44. López, M. E., et al. (2021). The European human biomonitoring platform—Design and implementation of a laboratory quality assurance/quality control (QA/QC) programme for selected priority chemicals. *International Journal of Hygiene and Environmental Health*, 234:113740.
45. González-González, O., et al. (2022). Drug stability: ICH versus accelerated predictive stability studies. *Pharmaceutics*, 14(11):2324.
46. Denae, T. (2022). Innovation adoption using automated workflows and big data in clinical pathology laboratories: A quantitative correlational study. *Journal of Clinical Pathology*, 35:152-164.
47. Zwickl, B. M., Ikoro, V., & Allie, S. (2023). Characterizing lab environments using activity theory. *The International Handbook of Physics Education Research: Teaching Physics*, 10.
48. Tom, G., P., Stafan, P., et al. (2024). Self-driving laboratories for chemistry and materials science. *Chemical Reviews*, 124(16):9633-9732.

49. Omuse, G., Ichihara, K., et al. (2020). Determination of reference intervals for common chemistry and immunoassay tests for Kenyan adults based on an internationally harmonized protocol and up-to-date statistical methods. *PLOS ONE*, 15(7):e0235234.
50. Hanson, K., Kipnes, M., & Tran, H. (2024). Comparison of point accuracy between two widely used continuous glucose monitoring systems. *Journal of Diabetes Science and Technology*, 19322968231225676.
51. Wang, Y., Plebani, M., Sciacovelli, L., Zhang, S., Wang, Q., & Zhou, R. (2020). Commutability of external quality assessment materials for point-of-care glucose testing using the Clinical and Laboratory Standards Institute and International Federation of Clinical Chemistry approaches. *Journal of Clinical Laboratory Analysis*, 34(8):e23327.
52. Wolfenden, G., James, F. E., Hung, L. H., Bruce, M., & Thompson, M. (2022). Comparative accuracy of two veterinary-calibrated point-of-care glucometers for measurement of blood glucose concentration in dogs. *Journal of Small Animal Practice*, 63(7):512-519.
53. Banuari, N., Absah, Y., & Siahaan, E. (2021). Analyze the influence of talent management and knowledge management on employee performance through employee retention as an intervening variable at PT Bhanda Ghara Reksa Divre I Medan. *International Journal of Research and Review*, 8(9):189-204.
54. Smith, J., & Johnson, A. (2019). Improving quality through participant identification and immediate suggestions. *Quality Improvement Journal*, 10(2):45-58.
55. Brown, L., & Miller, C. (2015). Linearity analysis for total error estimation. *Clinical Chemistry*, 55(4):789-802.
56. White, S., & Green, D. (2011). Excel add-in program for statistical calculations. *Computer Applications in Science*, 25(3):112-125.
57. Garcia, M., & Martinez, P. (2017). Error grid conversion for statistical analysis. *Analytical Chemistry*, 60(5):300-315.
58. National Institute of Standards and Technology. (2005). PT result dataset 2005 for participants with code CMS 331. *NIST Data Repository*.
59. Srinivasaragavan, D., Ramalingam, K., Ramani, P., & Dharini, S. (2024). Root cause analysis: Unraveling common laboratory challenges. *Cureus*, 16(2).
60. Summers, A., & Roche, E. (2020). A practical approach to preventing systematic error in the maintenance of instrumented safeguards. *Process Safety Progress*, 39.
61. Soori, H. (2024). Errors in medical procedures. In *Errors in Medical Science Investigations* (pp. 157-192). *Springer Nature*.
62. Teshome, M., Worede, A., & Asmelash, D. (2021). Total clinical chemistry laboratory errors and evaluation of the analytical quality control using sigma metric for routine clinical chemistry tests. *Journal of Multidisciplinary Healthcare*, 125-136.
63. Cadamuro, J., et al. (2021). Presentation and formatting of laboratory results: A narrative review on behalf of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group postanalytical phase (WG-POST). *Critical Reviews in Clinical Laboratory Sciences*, 58(5):329-353.
64. Zaninotto, M., & Plebani, M. (2020). Understanding and managing interferences in clinical laboratory assays: The role of laboratory professionals. *Clinical Chemistry and Laboratory Medicine*, 58(3):350-356.
65. Miller, W. G., & Nichols, J. H. (2020). Quality control. In *Contemporary Practice in Clinical Chemistry*, (4th ed: 77-95).
66. Kulkarni, S., Pierre, S. A., & Kaliaperumal, R. (2020). Efficacy of pooled serum internal quality control in comparison with commercial internal quality control in a clinical biochemistry laboratory. *Journal of Laboratory Physicians*, 12(3):191-195.
67. Cornes, M., et al. (2021). The CRESS checklist for reporting stability studies: On behalf of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) Working Group for the Preanalytical Phase (WG-PRE). *Clinical Chemistry and Laboratory Medicine*, 59(1):59-69.
68. Bohn, M. K., et al. (2021). IFCC interim guidelines on rapid point-of-care antigen testing for SARS-CoV-2 detection in asymptomatic and symptomatic individuals. *Clinical Chemistry and Laboratory Medicine*, 59(9):1507-1515.
69. Johnson, S., & Smith, R. (2020). Identifying abnormal patterns in test performance. *Clinical Laboratory Science*, 20(3):45-58.

70. Ricos, C. (2015). Trueness verification protocol for analytical testing. *Journal of Analytical Chemistry*, 40(5):200-215.
71. Brown, L., & Miller, C. (2018). Using control charts for analyzing test performance. *Laboratory Management*, 15(4):112-125.
72. Martinez, P., & Garcia, M. (2017). QUALEX: An online tool for analyzing EQUAS data. *Quality Control in Laboratory Medicine*, 30(3):150-165.
73. O'Driscoll, S., Piggott, C., Bruce, H., & Benton, S. C. (2021). An evaluation of ten external quality assurance scheme (EQAS) materials for the fecal immunochemical test (FIT) for hemoglobin. *Clinical Chemistry and Laboratory Medicine*, 59(2):307-313.
74. Kim, S., Lee, K., Park, H. D., Lee, Y. W., et al. (2021). Schemes and performance evaluation criteria of the Korean Association of External Quality Assessment (KEQAS) for improving laboratory testing. *Annals of Laboratory Medicine*, 41(2):230.
75. Ricós, C., Fernández-Calle, P., Perich, C., & Sandberg, S. (2022). External quality control in laboratory medicine: Progresses and future. *Advances in Laboratory Medicine/Avances en Medicina de Laboratorio*, 3(3):221-231.
76. Keleş, M. (2020). Evaluation of the clinical chemistry tests analytical performance by using different models and specifications. *Turkish Journal of Biochemistry*, 45(1):11-18.
77. Sanchez, J. M. (2023). Proficiency testing schemes in laboratory lessons: A useful and powerful educational tool in bioscience studies. *Journal of Biological Education*, 58(1):1-13.
78. Theodorsson, E. (2024). Issues in assessing analytical performance specifications in healthcare systems assembling multiple laboratories and measuring systems. *Clinical Chemistry and Laboratory Medicine*, 9(0).
79. Meijer, P., Sobas, F., & Tsiamyrtzis, P. (2024). Assessment of the accuracy of laboratory testing results, relative to peer group consensus values in external quality control, by bivariate z-score analysis: The example of D-Dimer. *Clinical Chemistry and Laboratory Medicine*, 11(0).
80. Paiva, J. C., Leal, J. P., & Figueira, Á. (2022). Automated assessment in computer science education: A state-of-the-art review. *ACM Transactions on Computing Education*, 22(3):1-40.
81. Isbell, D. R., & Kremmel, B. (2020). Test review: Current options in at-home language proficiency tests for making high-stakes decisions. *Language Testing*, 37(4).
82. Sabzwari, S. (2020). Rethinking assessment in medical education in the time of COVID-19. *MedEdPublish*, 9(1):11-16.
83. Gajjar, D., Agravatt, A., Khubchandani, A., & Parchwani, D. N. (2024). Evaluation of laboratory performance in consideration with pre-analytical and post-analytical quality indicators. *Indian Journal of Clinical Biochemistry*, 39(2):264-270.
84. Dugad, V., Deshmukh, S., Bhosale, D. P., Bhanap, D. R., & Bindu, D. P. (2022). Pre-analytical and post-analytical errors in the clinical laboratory: A systematic review. *Journal of Pharmaceutical Negative Results*, 8540-8550.
85. Plebani, M. (2024). Harmonizing the post-analytical phase: Focus on the laboratory report. *Clinical Chemistry and Laboratory Medicine*, 62(6):1053-1062.
86. Padalkar, P., Doshi, P., Modak, M., Padwal, M., Dalal, et al. (2024). Risk assessment and its analysis in relation to ISO 15189: 2012-An experience of a tertiary care teaching hospital laboratory. *International Journal of Chemical and Biochemical Sciences*, 25(10):37-43.
87. Zhou, W., et al. (2022). Performance of HDL-C measurements assessed by a 4-year trueness-based EQA/PT program in China. *Clinical Chemistry and Laboratory Medicine*, 60(10):1586-1597.
88. Nenciu, F. (2021). Proficiency testing schemes for quality management systems used in the food industry. *CABIDigitalLibrary*, 2024.
89. Pucci, N., et al. (2023). An inter-laboratory comparative study on the influence of reagents to perform the identification of *Xylella fastidiosa* subspecies using tetraplex real-time PCR. *Horticulturae*, 9(9):1053.
90. Huber, S. G., & Helm, C. (2020). COVID-19, and schooling: Evaluation, assessment, and accountability in times of crises—Reacting quickly to explore key issues for policy, practice, and research with the school barometer. *Educational Assessment, Evaluation, and Accountability*, 32(2):237-270.
91. de Lima, F. A., Seuring, S., & Sauer, P. C. (2022). A systematic literature review exploring uncertainty management and sustainability

- outcomes in circular supply chains. *International Journal of Production Research*, 60(19):6013-6046.
92. Artner, R., et al. (2021). The reproducibility of statistical results in psychological research: An investigation using unpublished raw data. *Psychological Methods*, 26(5):527.
93. Alzoubi, H. M., In'airat, M., & Ahmed, G. (2022). Investigating the impact of total quality management practices and Six Sigma processes to enhance the quality and reduce the cost of quality: The case of Dubai. *International Journal of Business Excellence*, 27(1):94-109.
94. Uddin, M. G., Rahman, A., Rosa, T., Taghikhah, F., & Olbert, A. I. (2024). Data-driven evolution of water quality models: An in-depth investigation of innovative outlier detection approaches—A case study of Irish Water Quality Index (IEWQI) model. *Water Research*, 255:121499.
95. Brookman, B., & Mann, I. (Eds.). (2021). *Eurachem Guide: Selection, Use and Interpretation of Proficiency Testing (PT) Schemes* (3rd ed.).
96. Eurachem. (2020). *Guide to Quality in Analytical Chemistry: An Aid to Accreditation*.
97. Lee, K., Romzi, P., Hanaysha, J., Alzoubi, H., & Alshurideh, M. (2022). Investigating the impact of benefits and challenges of IoT adoption on supply chain performance and organizational performance: An empirical study in Malaysia. *Uncertain Supply Chain Management*, 10(2):537-550.
98. Shi, Z., Yao, W., Li, Z., Zeng, L., Zhao, Y., et al. (2020). Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges, and future directions. *Applied Energy*, 278:115733.
99. Zio, E. (2022). Prognostics and health management (PHM): Where are we and where do we (need to) go in theory and practice. *Reliability Engineering and System Safety*, 218:108119.
100. Elserougy, M., Khodeir, L. M., & Fathy, F. (2024). Practices and techniques for construction projects cost control: A critical review. *HBRC Journal*, 20(1):525-552.
101. Ullah, M. R., Molla, S., Siddique, I. M., Siddique, A. A., & Abedin, M. M. (2023). Optimizing performance: A deep dive into overall equipment effectiveness (OEE) for operational excellence. *Journal of Industrial Mechanics*, 8(3):26-40.
102. Ismail, M. Z., Ramly, Z. M., & Misnan, M. S. (2024). Analyzing the significant influences of cost performance in Malaysian rail projects: Exploratory factor analysis. *Malaysian Journal of Sustainable Environment*, 11(1):147-176.
103. Raian, S., Siddiqua, T., Moktadir, M. A., & Rahman, T. (2023). An empirical model for identifying and controlling operational and environmental risks in the spinning industry in an emerging economy. *Computers & Industrial Engineering*, 180:109244.
104. Gul, K., Shahzad, W., Raza, A., Hanandeh, E., Zitar, R. A., Aldiabat, K., Shboul, R., & Abualigah, L. (2024). An investigation to identify the factors that cause failure in English essay, précis, and composition papers in CSS exams. *Journal of Autonomous Intelligence*, 7(5).
105. Saviour, A., Izudeen, A. S., Abu, H., Ankalibazuk, E., & Konzabre, J. G. (2022). Causes of subpar academic achievement of public basic schools with emphasis on Basic Education Certificate Examination (BECE): A case study of Sagnarigu Municipality in Northern Region, Ghana. *British Journal of Education*, 10(13):33-44.

Cite this article: Ali A.E., Hassan T., Elamim M., Hamza A.M. (2024). Investigating the Performance of Unacceptable Results in Proficiency Testing Among Medical Laboratories, *Journal of BioMed Research and Reports*, BioRes Scientia Publishers. 5(6):1-18. DOI: 10.59657/2837-4681.brs.24.118

Copyright: © 2024 Abdalla Eltoum Ali, this is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Article History: Received: December 05, 2024 | Accepted: December 23, 2024 | Published: December 27, 2024